Vol. 36, No. 6 November, 1929

PSYCHOLOGICAL REVIEW PUBLICATIONS

Psychological Review

EDITED BY

HOWARD C. WARREN, Princeton University
MADISON BENTLEY, Cornell University (*J. of Exper. Psychol.*)
S. W. FERNBERGER, Univ. of Pennsylvania (*Bulletin*)
W. S. HUNTER, Clark University (*Index*)
RAYMOND DODGE, Yale University (*Monographs*)
HERBERT S. LANGFELD, Princeton University, *Business Editor*

ADVISORY EDITORS

R. P. ANGIER, Yale University; MARY W. CALKINS, Wellesley College; JOSEPH JASTROW, University of Wisconsin; C. H. JUDD, University of Chicago; ADOLF MEYER, Johns Hopkins University; W. B. PILLSBURY, Univ. of Michigan; C. E. SEASHORE, University of Iowa; G. M. STRATTON, Univ. of California; MARGARET F. WASHBURN, Vassar College; JOHN B. WATSON, New York; R. S. WOODWORTH, Columbia University.

CONTENTS

The Measurement of Infant Behavior: Albert P. Weiss, 453.

General Affective Value: J. G. Beebe-Center, 472.

The Conditioned Reflex and the Sign Function in Learning: Katherine Adams Williams, 481.

A Functional Interpretation of the Conditioned Reflex: Clark L. Hull, 498.

Beats and Related Phenomena resulting from the Simultaneous Sounding of Two Tones—II: Ernest Glen Wever, 512.

Critique of the Seashore Consonance Test: A Reply to Dr. Larson: Christian Paul Heinlein, 524.

Discussion:

Standardizing the Grading of Laboratory Reports: H. D. Marsh, 543.

Purposive Behavior and Prepotent Stimulus: Zing Yang Kuo, 547.

Note on Use of the Term 'Observer': J. F. Dashiell, 550.

PUBLISHED BI MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Psychological Review Publications

OF THE

American Psychological Association

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
RAYMOND DODGE, YALE UNIVERSITY (*Monographs*)
SAMUEL W. FERNBERGER, UNIV. OF PENNSYLVANIA (*Bulletin*)
W. S. HUNTER, CLARK UNIVERSITY (*Index*)
MADISON BENTLEY, CORNELL UNIVERSITY (*J. of Exp. Psych.*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, *Business Editor*

WITH THE COOPERATION OF
MANY DISTINGUISHED PSYCHOLOGISTS

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bimonthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 500 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly, the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bimonthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 500 pages.

PSYCHOLOGICAL INDEX

Is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL MONOGRAPHS

consists of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

ANNUAL SUBSCRIPTION RATES

Review: $5.50 (Foreign, $5.75)
Journal: $6.00 (Foreign, $6.25)
Bulletin: $6.00 (Foreign, $6.25)
Monographs: $6.00 per volume (Foreign, $6.30)
Psychological Index: $4.00

Current Numbers: Review or Journal, $1.00; Bulletin, $.60

Review and Bulletin, $10.00 (Foreign, $10.50)
Journal and Bulletin, $11.00 (Foreign, $11.50)
Review and Journal, $10.00 (Foreign, $10.50)
Review, Bulletin, and Journal, $15.00 (Foreign, $15.75)
Review, Bulletin, Journal, and Index, $18.00 (Foreign, $18.75)

Subscriptions, orders, and business communications may be sent direct to
PSYCHOLOGICAL REVIEW COMPANY,
PRINCETON, N. J.

PSYCHOLOGICAL REVIEW PUBLICATIONS

Psychological Review

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY
MADISON BENTLEY, CORNELL UNIVERSITY (*J. of Exper. Psychol.*)
S. W. FERNBERGER, UNIV. OF PENNSYLVANIA (*Bulletin*)
W. S. HUNTER, CLARK UNIVERSITY (*Index*)
RAYMOND DODGE, YALE UNIVERSITY (*Monographs*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, *Business Editor*

ADVISORY EDITORS

R. P. ANGIER, YALE UNIVERSITY; MARY W. CALKINS, WELLESLEY COLLEGE; JOSEPH JASTROW, UNIVERSITY OF WISCONSIN; C. H. JUDD, UNIVERSITY OF CHICAGO; ADOLF MEYER, JOHNS HOPKINS UNIVERSITY; W. B. PILLSBURY, UNIV. OF MICHIGAN; C. E. SEASHORE, UNIVERSITY OF IOWA; G. M. STRATTON, UNIV. OF CALIFORNIA; MARGARET F. WASHBURN, VASSAR COLLEGE; JOHN B. WATSON, NEW YORK; R. S. WOODWORTH, COLUMBIA UNIVERSITY.

VOLUME XXXVI, 1929

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

LANCASTER PRESS, INC.
LANCASTER, PA.

CONTENTS OF VOLUME XXXVI

BEEBE-CENTER, J. G., General Affective Value 472–480
BORING, E. G., The Psychology of Controversy 97–121
BROWN, J. F., The Methods of Kurt Lewin in the Psychology of Action and Affection 200–221
CALKINS, M. W., Analysis: Chemical or Psychological? 348–352
DASHIELL, J. F., Note on Use of the Term 'Observer' 550–551
DAVIS, R. A., & GOULD, S. E., Changing Tendencies in General Psychology 320–331
DIMMICK, F. L., A Reinterpretation of the Color-Pyramid 83– 90
FEARING, F., René Descartes: A Study in the History of the Theories of Reflex Action 375–388
FLETCHER, J. M., An Old Solution of the New Problem of Instinct 44– 55
GEISSLER, L. R., The Objectives of Objective Psychology 353–374
GESELL, A., Maturation and Infant Behavior Pattern 307–319
GUNDLACH, R. H., Four Sources of Confusion in Psychological Theorizing .. 285–306
HARRIS, A. J., Theoretical Studies from the Harvard Psychological Laboratory: Analysis: A Contribution to Psychological Method 1– 12
HEINLEIN, C. P., Critique of the Seashore Consonance Test: A Reply to Dr. Larson 524–542
HOLMES, S. J., The Freaks of Creative Fancy 446–449
HULL, C. L., A Functional Interpretation of the Conditioned Reflex 498–511
HUTCHINS, R. M., & SLESINGER, D., Legal Psychology 13– 26
JOHNSON, H. M., Did Fechner Measure 'Introspectional' Sensations? 257–284
KUO, Z. Y., Purposive Behavior and Prepotent Stimulus 547–550
KUO, Z. Y., The Net Result of the Anti-Heredity Movement in Psychology . 181–199
LOUTTIT, C. M., The Use of Bibliographies in Psychology 341–347
MARSH, H. D., Standardizing the Grading of Laboratory Reports 543–547
MILES, W. R., Horizontal Eye Movements at the Onset of Sleep 122–141
MINER, J. B., The Procedure of Thinking about Mind 332–340
PRATT, C. C., Theoretical Studies from the Harvard Psychological Laboratory: Faculty Psychology 142–171
ROBERTS, W. H., A Note on Anthropomorphism 95– 96
ROSENOW, C., In Reply to the Rescuer 450–451
RUCKMICK, C. A., A New Classification of Tonal Qualities 172–180
STRATTON, G. M., Emotion and the Incidence of Disease: The Influence of the Number of Diseases, and of the Age at Which They Occur 242–253
SWIFT, E. J., The Learning Process: A Criticism and a Theory 27– 43
THURSTONE, L. L., Theory of Attitude Measurement 222–241
TRYON, R. C., The Interpretation of the Correlation Coefficient 419–445
WEISS, A. P., Some Succor for Professor Kuo 254–255
WEISS, A. P., The Measurement of Infant Behavior 453–471
WEVER, E. G., Beats and Related Phenomena resulting from the Simultaneous Sounding of Two Tones 402–418, 512–523
WILLIAMS, K. A., The Conditioned Reflex and the Sign Function in Learning 481–497
WINSOR, A. L., Inhibition and Learning 389–401
WITTY, P. A., & LEHMAN, H. C., Further Remarks upon Character Testing . 91– 95
WITTY, P. A., & LEHMAN, H. C., Religious Leadership and Stability 56– 82

VOL. 36, No. 6 November, 1929

THE PSYCHOLOGICAL REVIEW

THE MEASUREMENT OF INFANT BEHAVIOR[1]

BY ALBERT P. WEISS

Ohio State University

The Difference between the Human and the Animal Infant.—At birth the human infant is a biological organism. With growth, coördinated activities develop. The arms and legs move more rapidly, and many of the movements are repeated. The change in environment with birth introduces an entirely different set of stimuli than are active before the infant is born. The handling of the new-born infant by the physician, nurse, and mother introduces a social factor. The instruction of doctor to nurse and mother as to the care of the child depend on his training as a physician. The way in which the mother handles the infant after she leaves the hospital will depend upon her physician, mother, and friends, plus whatever she herself may add.

All of this socialized stimulation or social inheritance is added to the physical inheritance of the infant. This social form of stimulation is not present for the animal young, and this makes the study of the behavior of the human infant something different from the study of the animal infant. The sense organs of the human infant are fairly well developed, the central nervous system not so well developed, but the

[1] Presidential address before Midwestern Psychological Association, University of Illinois, May 12, 1929.

The experimental results upon which this paper is based will appear as a monograph in the near future, under the joint authorship of Dr. K. C. Pratt, Dr. Amalia K. Nelson, and Dr. K. H. Sun. The experiments represent a period of four years. The present paper presents some of the theoretical principles which were developed during the earlier investigations and the investigation which is at present going on.

muscles and glands seem to be functionally complete. In the animal young the coördination of the skeletal musculature is more highly developed.

Relation between Infant and Adult Behavior.—To describe infant behavior it is necessary to know what to expect. This can only come through observation. We must learn that new-born infant behavior is not a faint copy of adult behavior, and must understand that the infant is fairly well adapted to his own environment for his particular age, even though his movements do not resemble the adult's movements in the same physical environment. When we refer to adult behavior, we mention such activities as walking, peeling an orange, getting Johnny ready for school, paying taxes, giving a lecture, playing golf. Infants' behavior cannot be described by these classes. However, we do know that the adult activities replace the infantile reactions, and one of the problems of psychology is to learn how this change takes place. When we observe the movements that infants make, we find that there may be a movement of the head, of the right arm, of fingers or toes, the trunk may turn slightly the eyes may open or close, all of this in a very irregular manner. As the feeding period approaches, these movements become more frequent, and a number of them occur at the same time. There is very little coördination of the different movements. While each particular movement is probably a reaction to some specific stimulus, it is practically impossible to designate the sense organs that are being stimulated. The assumption that no movement is ever *repeated exactly* becomes a very obvious fact. It seems at first futile to speak of classifying the behavior. While this is true theoretically, as a matter of fact we do group together certain types of movements according to their resemblances. This leads to much overlapping and forms one of the uncertainties in the description of behavior.

The existing classifications or dictionary definitions are of little help. Thus the term 'crying' as defined in the 'Standard Dictionary' is given as "To utter or announce loudly and publicly; shout out; proclaim specifically; to

advertise with the voice; to beg or beseech for; implore; to call; to utter in a loud or vehement tone a sound whether articulate or inarticulate; to clamor; shout; yelp; howl; lament."

These represent only a few of the definitions which seem appropriate for registration, yet we can see that they do not meet our requirements. None of the definitions is clearer than the original word 'crying.' For scientific purposes, we need some form of registration or description which can be made uniform from one experimenter to another without too much training. Ideally we would meet this situation by being able to register the amount and extent of the movements of the muscles involved. From moving pictures we could demonstrate such activities as squirming, sucking, smiling, etc. In crying, for instance, we might register the actual sounds. From such an analysis it would be possible to classify the different kinds of cries. For most studies the phono-photographic method of recording is too costly or too complicated. Auxiliary recorders such as tambours, voice keys, pneumographs and other attachments change the stimulating conditions so much that the reactions are distorted. For those who have worked with infants it is clear that it will be necessary to design a special technique.

When does Infant Behavior Begin?—The actual duration of infancy should be agreed upon. Medical practitioners vary. For some, infancy includes the day of birth and the first three months. For others infancy begins with the day of birth and ends when the baby is weaned. Of course when the actual age of the infant is reported, it makes very little difference, but in the literature the age is so frequently omitted that it is difficult to make comparable observations. However, with the more careful experimental work which is now being done this difficulty will vanish.

When we ask the question, 'When shall we begin to study infant behavior?' the very obvious answer is, the earlier the better. However, scientists are more or less restricted by the experimental technique they have acquired, and for the study of pre-natal stages a technique and method is used that

is very different from that used in the study of the child after birth. Birth, then, marks a period when the method of investigation changes. There is of course always the tendency to push the researches to earlier and earlier stages. This is the case with all science. At Ohio State we are now studying new-born infants from one or more hours old up to twelve days. We have called this the period of the new-born. It is rather sharply marked off from earlier or later stages. The day of birth presents a decided change in environmental conditions. During the first week there is an adaptation to this change, and by the end of the second week most infants have established rather regular feeding habits, and are gaining in weight. This first two-weeks period, then, we believe might well be called the period of the new-born infant.

What Can One Study about Infant Behavior?—The lack of coördination in the new-born infant is gradually replaced so that in the adult we have such highly specialized motor coördinations as are found in the handicrafts and in the professions and in that form of behavior called thinking. The crying, squirming infant becomes the motor-driving, golf-playing leader of men whose verbal reactions as stimuli for others may determine the destiny of a nation. What are the steps from the practically helpless infant to those finely coördinated movements in the personal, domestic, public, vocational, and recreational behavior of the adult? We do not know how the transition from one form of activity to the next is made, nor have we a terminology which is adequate to describe the development of the life career of a given individual.

From the scientific standpoint this means that we must start with the assumption that at birth the organism is only capable of making movements. Further, these movements are not very well coördinated in the sense that only specific movements follow upon specific forms of stimulation. We can almost say that when the stimuli are not too strong, any movement may follow upon any stimulus. The coördination which slowly develops is the product largely of the social

stimuli which are operative on the child, and these social stimuli are not the development of this generation, but have a historical background which it is practically impossible to trace backward to its origin. We are beginning to realize that perhaps it is not necessary to postulate a mind which takes over the responsibilities of forming those habits which make up adult behavior. But when we drop the mind, our problem is expanded. We must study just what reactions follow upon certain forms of biophysical and biosocial stimuli, and the extent to which the reactions of the infant itself are biosocial stimuli for the individuals with whom it is in direct contact. The child modifies his social environment almost as much as he himself is modified. Certainly this is true of his mother and the family, and to a lesser extent of his grandparents, aunts, uncles, nieces, and the neighbors.

At Ohio State we are working with infants during the first two weeks of life. Our problem is predominantly one of controlling the biophysical stimuli or trying to determine what they are. For the most part the social environment of the infant in our work is the experimenter who is conducting the experiment.

The reactions which we are investigating are the movements during the control period, reactions to light, sound, taste, smell, temperature, holding nose and arms, the plantar reaction, and the sucking reaction. At present Mr. O. C. Irwin is making an intensive study of the behavior of infants during the whole twenty-four hours and for the time they are at the hospital. We are trying to hold the external stimulating conditions constant. The infant is under continuous observation except for those periods which are spent with the mother and during which he is bathed and fed. Such a schedule requires a number of trained observers, and we have been obliged to call in graduate students. This introduces a variable factor in the observations. Not all observers report in the same way. However, from our data we shall be able to calculate the degree of variability. It would be more scientific, of course, to improve the observational technique, but so many observers are required that

it is practically impossible to secure them. To watch an infant through a peep-hole and record all the movements while the infant is screaming at the top of his lungs for one and sometimes two hours is grilling work. We have run a check on the variability between experimenters by having one of them (Mr. Irwin) report all the movements during the longest interval between feeding periods, which extends from about ten o'clock at night to about six in the morning.

Conditions under which Observations are Made.—The conditions in the nursery vary greatly in their stimulating effects upon the infant. The illumination is constantly changing, sounds occur, there are uncontrolled temperature and tactual stimulations, and the nurses handle the babies. To secure more uniform stimulating conditions we have a separate room in the maternity ward. In this room we have a cabinet large enough to conveniently hold the infant and such accessories as are necessary for experimentation. This cabinet is illuminated only from within. It is relatively soundproof, has regulated temperature and humidity, and is equipped with a stabilimeter platform upon which the infant is bedded. This stabilimeter records the gross bodily movements on a moving strip of paper. On the same strip there is also a time marker and a marker for the various stimuli which are used.

The infant is brought from the nursery in its bassinet. Its own blanket is placed on the stabilimeter platform. All clothing is loosened up so that the movements are free. No covering is placed over the infant's body, and the temperature is kept between 80° and 85° F. Humidity readings are made and the light reduced to a point where the experimenter is just able to discriminate the movements of the infant. The cabinet is then closed and the infant observed by one or both of the experimenters through peep-holes. The experimenter at the registering apparatus records the specific movements of the infant on the paper record.

The Control Period.—When a stimulus is applied and the infant reacts, a number of questions arise. Would the infant have reacted in this way without the stimulus? How frequently does the infant react to this stimulus in this particular

manner? Does the reaction change with age? Does it vary for different infants, etc.? Even without careful observation, one soon finds that some infants are more active than others. Some vary more from one hour to the next than others do throughout the day. Obviously, comparing the reactions of two extreme infants for the same experimental stimulus conditions when the uncontrolled activities of the infants show wide variations, reduces the reliability of whatever data are secured.

In order to make comparisons we soon realized that it was necessary to have some kind of a control. We decided to use a period during which the external stimuli were reduced to a minimum. There is of course no control over the internal stimulations, which form a very important factor in infant behavior.

Method of Recording Movements.—The experimenter records the movements by a series of symbols which indicate the specific movements as they occur. These symbols cover about 200 movements and conditions and are of the following order: *aw* for awake, *dr* for dry; *P3*, lying on left side; *P7*, lying on back facing up; *F9*, extension of toes; *B10*, back arched; *S10*, sucking; *S13*, audible breathing, etc. The plan of the symbols is more logical than appears from these samples, so that each experimenter can easily memorize them. Greater objectivity is secured than when a given reaction is defined. There is also a simplification in the subsequent organization of the movements into various categories of segmental activities such as a combination into eye movements, head movements, facial movements, sounds, body movements, etc. During the control period or those periods in which no stimuli are given, one experimenter is sufficient, and he observes the infant through a peep-hole in the cabinet. Whenever the reactions are to controlled stimuli, there are two experimenters, one who administers the stimuli, and another who records the movements.

At intervals the experimenters check up on each other to determine whether they are recording uniformly. This plan of double observation and recording is essential for observing

reactions to controlled stimuli. It is impossible for one person to note and record all that is going on. Photographic moving pictures would solve this difficulty, and we tried them, but the expense is great. The tedious process of making motion-studies after the film has been secured, the strong illumination, the noise of the lamp—all these factors seem to give less usable results than the simpler method of recording the movements by symbols on the stabilimeter record.

The question arises, why try to make such a careful description of the overt movements? Is not the energy of the activity itself a sufficient index of the behavior? The difficulty lies in the fact that the energy record alone does not measure the degree of motor coördination. These coördinations may not be accompanied by any marked energy change. The rate of change in coördination is not the same for all infants nor for the same infant at different times, but the object in studying the infants at this early age is precisely that of trying to trace the increase in motor coördination.

The qualitative record of the movements is significant in studying the genetic development of the infant. Granting its unsatisfactory character we are still able to make behavior differentiations which seem significant. I mention the fact here to call attention to the need for a syllabus and set of symbols for describing infant behavior. Our own problem with new-born infants is rather simple, but when the child grows older, the problem becomes more complex, and the whole method of recording infant and child behavior is still to be developed.

We are trying to anticipate the problem of recording the activity of older infants. We have about reached the stage where the stabilimeter can no longer be used. A sturdy infant at ten days now almost taxes its capacity, and while of course it could be made larger and stronger, we are in doubt as to what it will show under stimulating conditions. The activity record is valuable for measuring the control period in which the reactions are largely due to internal stimulations, but when we try to measure the effects of different kinds of stimuli, such as light compared with sound, smell compared

with temperature, etc., the activity records become less valuable. The bodily reactions to light are not much stronger than the bodily reactions to sound, and they may be less. As we see the problem at present, it will not be one of recording the intensity of the movements as much as recording the kinds of movements.

Activity Measurements.—The stabilimeter is a sensitive platform upon which the infant lies. Any movements that occur are recorded by two pens. One of these records the movements along the head-foot axis, the other along the right-left axis. Every movement, no matter in what direction, is thus expressed in two components. This we have called our activity record. Actually, it consists of two wavy lines. In measuring the activity we count the number of oscillations of each of the pens and the amplitudes of the vibrations in millimeters. Our record is thus movement in millimeters of the stabilimeter pens. The question will arise, how closely does the stabilimeter record correspond to the actual movement of the infant? The answer is that it is a relative measure. A movement of 10 millimeters of the pens does not mean that the infant moved 10 millimeters. He may have moved suddenly with only a leg, or he may have moved very slowly with the whole body, and in both cases the record might show 10 millimeters. We can only say that in the long run there is a close correspondence between the stabilimeter readings and the actual energy expended by the infant. This, however, is a conclusion based only on superficial observation. Much more work is necessary before the actual expenditure of energy is recorded. An infant cannot be restricted by belts, head-pieces, tambours, etc., and give normal reactions. We assumed that the amount of stabilimeter movement was a fairly close function of the actual movement of the infant, and on this assumption then asked the following questions:

1. Shall we measure activity as gross activity, that is, as millimeters of movement between stimulations?

2. Shall we measure activity as specific activity, that is, measure only that activity which seems definitely a reaction

to the given stimulus, and leave out that activity which seems to be a secondary effect?

3. Shall we introduce a time unit into our measurements in the form of millimeters per stimulus per second?

4. Shall we introduce weight also, that is, millimeters per stimulation, per second, per kilogram of body weight?

Of course the only answer is to try out all of them and from the data try to determine which is the best measure. This we proceeded to do. We started with the control period and calculated the relationships between the following conditions:

1. Millimeters movement per kilogram.

2. Millimeters movement per kilogram per second.

3. Percentage of time infants were actually moving as compared to the total time of the control period.

For the experimental period, we added three other values:

4. Millimeters movement per stimulation.

5. Millimeters movement per stimulation expressed as a percentage of the total amount of movement for some given category of activities, as for instance the total amount of movement to the smell stimuli, taste stimuli, etc.

6. Number of stabilimeter oscillations per minute.

We have not yet reached a decision as to which of these measures is the best one. We have the data which will enable us to calculate the relative reliability of all the measures, but this problem is such an involved one that we decided upon some temporary expedient which seemed to be fairly reliable and at the same time, easy to understand. We selected millimeters of movement per stimulation, and percentage of movement per stimulation for a given class of stimuli. These measures cannot, of course, be applied to the control period, because there were no external stimulations, and it was impossible to determine what internal stimuli were operative, but the amount of movement per minute during the control period is a good reference point from which to compare the reactions to different stimuli. On account of the time involved in measuring both amplitude and frequency, we are now using No. 6, number of stabilimeter

oscillations per minute. When it is recalled that the record of one infant is over a mile in length and that during a considerable part of this length the oscillations and notations of movements are packed so closely that it is difficult to read them, the reader may form some idea of the task involved.

Duration of Control Period.—We tried to answer a number of other questions about the control period: (*a*) whether it should occur before, during, or after the experimental period, (*b*) how long it should last. We began by using a 15-minute control period at the end, but this soon proved unsatisfactory, because if an infant became unruly during the experimental period and had to be returned to the nursery, we lost what data we had secured.

We finally selected a control period of 10 minutes. Three forms of distribution are used: (*a*) 10 minutes at the beginning; (*b*) 10 minutes in the middle; (*c*) 5 minutes at the beginning and 5 minutes at the end. This avoided any time and space errors and made it unnecessary to keep track of the periods. For a given day, all the infants were given 10 minutes before experimenting; the next day all control periods were in the middle of the experimental series; the third day, all infants received a 5-minute pre-period and a 5-minute end-period. This simplified the schedule, and since each infant had on an average eight experimental periods, mere chance would distribute the periods fairly well. When it is impossible to foresee just what infants will be available, and where the individual records have not yet been worked into a form where back references are easy, it is better to rely upon chance.

The Advantages of Reporting Simple Movements.—When one reads of the reactions of infants in the literature, one is struck by the extent to which the experimenter reports infant behavior in terms of what is found in adults. For instance, in investigations of the Babinski reflex, now called the plantar reflex, the earlier reports concerned themselves with the flexion and extension of the toes. Later, it was noted that leg and hip movements also occurred. Then it was noted that the reflex did not occur at every stimulation

and that it occurred differently whether the experimenter held or did not hold the leg. The reports were gradually extended to indicate that sometimes the reflex did not occur upon stimulation of the sole of the foot, but sometimes occurred without the stimulation of the foot. Instead of a relatively simple reaction to a definite stimulus, the plantar reaction was found not to be a simple extension or flexion of the toes. It occurred rarely in isolation, and as a rule other bodily movements such as head movements, eye movements, movements of the arms, fingers, etc. also appeared. Furthermore, the early description which Babinski gave to this reflex so predisposed later investigators to look just for this particular kind of movement and no other that only those instances were reported where the stimulation definitely released the reflex. The cases in which the stroking of the foot did not result in flexion or extension of the toes were not reported, nor those in which toe flexion and extension occurred without plantar stimulation. The contradictory reports on infant behavior are largely due to this restricted type of observation. The experimenter sees what he has prepared himself to see and ignores what actually occurs but is not expected. By restricting the reports of behavior to actual movements of the bodily segments some of this prejudice is overcome. Of course for special types of stimulation it may be necessary to group particular sets of specific movements into a motor coördination group, simply for ease in reporting. For instance, the reaction to a weak sour stimulus applied to the tongue is very frequently a grimace, opening of the eyes if closed, or an increase in the lid opening if already open, and a kind of chirp which is very characteristic. Among ourselves we designated this as the 'sour' reaction, but without a previous analysis into the simpler movements, the designation 'sour reaction' would be ambiguous.

Crying or Sounds in Relation to Movement.—In the newborn infant, sounds or vocalizing seldom occur without bodily activity of some kind. Even the cry that precedes the violent hunger reactions does not begin until after the movements of bodily restlessness have started. This is an im-

portant theoretical consideration for the development of language. In a way, the cry is language. It is a signal which is very effective in stimulating adults. If, as seems to be the case, vocalizing develops along with bodily movement, there is an opportunity for understanding how language may develop. At first crying is only an accompaniment of violent bodily movement. Later sounds accompany not only violent movements, but movements which require less bodily energy. In this way, the sounds become more numerous and instead of the violent cry in which all sounds are uttered in all kinds of combinations, there is a breaking up of the sounds into more or less separated groups which are uniform for the type of movements that the infant is making.

As soon as this takes place, the mother uses these sounds as signals of what the infant is doing. This again emphasizes the biosocial character of the sounds. From our experiments, it seems that one of the important steps in infant investigations is to determine this relation between sounds and the corresponding bodily movements.

The Hunger Reaction.—In this reaction all movements are exaggerated, and many movements are going on at the same time. The infant is writhing, throwing himself about, kicking, slashing, howling, and expending an amount of energy which almost taxes the capacity of the recording device. Mr. Irwin has called these movements mass reactions. They are uncoördinated to the extent that they seem to be convulsive.

This hunger reaction may be regarded as the first reaction which the infant makes to its *social* environment. The infant is merely producing the maximum number of stimulations which are likely to act on some other individual. As specific behavior these movements are uncoördinated reactions, but from the biological standpoint, they are adaptive in that they stimulate someone else to do something for the child. How strong a stimulus these hunger reactions are for others is shown when we watch an experimenter who is carrying on his first observations. We find him almost as much disturbed as the infant. He watches the clock, makes

errors in his records, slips up on the experimental technique, drops of perspiration form, his movements become hurried and all his actions indicate that he is working at a much higher tonicity level than when the infant is sleeping. The calm and so-called impassive experimenter who is supposed to be merely recording the specific movements of the infant, can scarcely refrain from breaking his technique and trying to 'quiet' the infant in a truly motherly or fatherly sense. However, after a few feeding periods have been observed, the experimenter becomes more passive during the hunger reaction, especially when he finds that all the other infants in the nursery who are doing the same thing seem to survive.

It seems strange to say that a series of uncoördinated movements are the most adaptive form of behavior, yet such is the case in the ordinary home. Of course in the hospital the stimulus is not nearly so effective. The nurses in the maternity ward have learned that crying is a perfectly normal reaction when the infant is hungry, and they will go calmly about their duties with a dozen infants all going strong.

I have used the term convulsive reaction. This is not the correct term. The infant is screaming, kicking about, throwing himself from one side to the other, slashing with arms, jerking the legs, but any external stimulation such as flashing a light into its eyes, stroking the body, making a loud sound, has a very immediate pacifying effect. Crying stops, activity stops, the infant quiets down with a speed that is startling. If the most effective stimulus, that of placing the nipple in the infant's mouth, is used, the effect seems magical. This indicates that the hunger reaction is not a convulsion in the medical sense. From the social standpoint the hunger reaction produces a strong biosocial stimulus for others. I introduce this point at this time to indicate that while an activity may be biophysically uncoördinated, it is nevertheless a strong biosocial stimulus. This is significant in reporting infant behavior. Many reactions which would be reported as non-specific or useless movements so far as the infant is concerned may be very effective biosocial stimuli.

Infant behavior thus has a dual character:

(1) The biophysical activity which is measured by the degree of coördination between the muscles to produce some given effect such as fixating, reaching, grasping, etc.

(2) The extent to which the movements produce stimuli which act on other individuals so that they will react to the infant.

The Development of Skill.—During the 10-day period in which we have the infant, not much can be determined as to the development of specific reactions. The first week is virtually a period for the recovery from the shock of birth. The second week shows some advances in the direction of motor coördination, but the changes which simple differences of feeding produce tend to mask whatever coördination has developed. The experimenter himself can see that improvement is taking place in the greater precision of the movements and the time relations between them, but this is difficult to describe. There is less interference of movements of the infant's arms and legs. The differences are most striking when a new-born infant is compared directly with a 10-day infant. The older infant does about the same things as the new-born infant, but all movements are carried out more rapidly and with greater precision. The experimenter is able to foresee better what the older infant is going to do, and there is a closer relation between stimulus and reaction, even when the reactions do not occur any more frequently.

The Application of Recent Theories to Infant Behavior.—The recent appearance of Professor Köhler's book and the excellent description of the gestalt principles which it contains have encouraged me to compare the gestalt hypothesis with some of the theories held in this country. If we take such a reaction as sucking, for instance, there is no doubt that the infant even during the first day does react differently to the nipple than to other stimuli. The difference is not very great, and the reactions may be wrong at first, but the speed with which the infant learns to nurse indicates that the reaction is already organized in the gestalt sense. If we take

the sucking reaction under experimental conditions, as when the child is just beginning to show the hunger reaction, and we are investigating the effect of water of a particular temperature introduced into the mouth, the following conditions appear:

The child is restless, moving arms, legs, head, body, with perhaps some throat movements. If the mouth is closed, the experimenter touches the upper lip lightly, and the mouth opens. At the same time a few drops of water at a known temperature are squirted into the child's mouth, and if the temperature is in the neighborhood of 40° Centigrade, the infant quiets down immediately. From the standpoint of gestalt we may say that the infant has already discriminated the warm water stimulation out of the total field. We will not raise the question as to whether the child has *experienced* the warm water. Köhler himself usually handles this point by the phrase, "If the child has any experience at all, . . ." etc. Disregarding this, there is actually a very pronounced pacifying effect from the water stimulus, which as an energy unit is probably much less than the totality of the hunger stimuli which come from the internal sense organs. The gestalters would say that the warm stimulus on the tongue organizes the total field into a fore- and a background, with the hunger stimulations forming the background against the experience of warm water. In trying to account for this Köhler would limit his description to what happens in the psychophysical area of the brain. The actual processes involved, Köhler calls the dynamo-geometrical relations of these two processes in the brain.

As a behaviorist, I would include more of the reaction mechanism than Köhler includes. I would say that the warm water pacifies so quickly because the hunger reaction is not only limited to crying and throwing the body about, but that some of the processes of the hunger reactions increase the sensitivity of the sense organs. This assumption is supported by the following evidence. When the child is reacting strongly to hunger stimuli, there are also incipient mouth movements such as pursing the lips, moving the

tongue, and assuming a posture of the head (face up) so that any stimulus which may act upon the mouth will have favorable conditions for producing a maximum effect. In this sense the hunger reaction prepares for the form of stimulation which will stop the reaction itself. Now whether the hunger reaction is such a finely coördinated series of movements that only the mouth will be sensitized is very doubtful. Taking up the child, fondling it or walking the floor with it also pacifies the infant, although this effect is temporary, because these stimuli do not remove the hunger stimuli such as the contractions of the stomach. It is only the circumstance of the biosocial relation between infant and mother that finally strengthens the sucking reaction under hunger stimulation. The mother has been taught by others that feeding the child is the most effective method of pacifying it and herself also.

From the analysis as given, it hardly seems necessary to speak of the hunger reaction as a drive. It starts very slowly and intermittently. If a food stimulus does not occur, the reaction increases in intensity. We cannot even say that the stimuli become stronger. We only observe that the hunger reactions become stronger.

In the terminology of the gestalt school there is a progressive organization of the total field into a background of organic experiences and a foreground of a warm, pleasant experience.

The point I wish to make is that to say that the hunger reaction has a drive to it, or that it has been organized into a gestalt does not add anything to what has already been reported. To add to the detailedness of the report it is necessary to make further observations and do more experimenting. The conceptions which are included under the term drive or gestalt are merely terms which serve as a guide in directing further research. There is no probability of finding a drive or a gestalt in the anatomical, physiological, or activity sense. In other words, they are hypothetical principles which have their influence in directing subsequent research. To use the hypothesis of a 'drive' leads to further

physiological researches; to use the hypothesis of gestalt leads to further 'experience' researches.

The Hypothesis of a Biosocial Component in Infant Behavior. —In the investigation of human behavior I have continually stressed the fact that there is a social factor in the development of human responses, and that if we begin to investigate this biosocial factor, many of the responses that human beings make will become clearer than if we attribute them to drives, gestalt, or mental factors. As a practical application of this biosocial hypothesis the work we have done with the hunger reaction thus far indicates that there is a social component, and the next immediate step is that of (1) observing the actual movements in greater detail, (2) isolating the internal stimuli for the crying reaction, and (3) finding a biosocial factor which modifies the crying into its adult form.

What does the infant do as soon as the mother takes it up? This is a whole series of problems in which the age of the child and his previous reactions to similar conditions must be considered. What does the mother do when she hears the child cry, and in what way does her handling of the child modify the crying reaction? The same thing may be said of the physician and even of the experimenter who is observing the child. We have found that we cannot ignore these external factors. The experimenter or the teacher is not merely a sort of recording machine that has no effect on the movements of the infants.

There is also the historical factor. The mother will react to the crying as she has been taught to react by the physician, nurse, or her own mother. This biosocial hypothesis also leads to investigations as to what mothers have done in the past when their babies cried, and in what way their actions have modified the crying or the behavior of infants. This leads to problems in anthropology which at present we have not even tried to solve. Further, by introducing the biosocial component we may find it unnecessary to use such troublesome terms as experience, consciousness, drive, organized whole, etc. Our investigations will proceed in the same way that observations in the other sciences proceed. The

hypothesis of a genetic as well as a biosocial environment actually extends our program of experimentation. In other words, by including the mother, the doctor, and the experimenter in with the infant in our investigations we learn not only more about infant behavior but more about adult behavior as well.

[MS. received May 31, 1929]

GENERAL AFFECTIVE VALUE[1]

THEORETICAL STUDIES FROM THE HARVARD PSYCHOLOGICAL LABORATORY

BY J. G. BEEBE-CENTER

Harvard University

The problem of assigning to an object a general affective value, one which is not dependent upon a particular observer, has long puzzled psychologists and æstheticians. The averaging of divergent judgments—and in affection individual differences are very marked—at once meets with the criticism that 'En fait de gôut, plus encore qu'ailleurs, l'élite prime la masse.'[2] Recourse to the elite, on the other hand, fails to be accepted because of the extra-psychological considerations involved in determining the elite. In the present paper I wish to propose a method for dealing with the problem which, though it avoids arbitrary democracy by differential weighting, derives the different weights directly from the affective data.

The method is essentially an application to the field of affection of the procedure proposed by Spearman for the determination of a general intelligence factor in mental tests. Spearman writes:

> Whenever the tetrad equation holds throughout any table of correlations, and *only* when it does so, then every individual measurement of every ability (or of any other variable that enters into the table) can be divided into two independent parts which possess the following momentous properties. The one part has been called the 'general factor' and denoted by the letter g; it is so named because, although varying freely from individual to individual, it remains the same for any one individual in respect of all the correlated abilities. The second part has been called the

[1] I wish to express my thanks to Dr. Donald Snedden for considerable assistance in connection with the statistical phase of this paper.

[2] C. Lalo, Notions d'esthétique, Paris, 1925, 22.

'specific factor' and denoted by the letter s. It not only varies from individual to individual, but even for any one individual from each ability to another.[3]

Spearman points out that the application of this theorem is quite general—indeed, he states that it is not even restricted to psychology. It is consequently perfectly legitimate to extend it from the field of mental testing to that of affection, provided only that it be possible to interpret the terms involved in the theorem in such a way as to fit the data of experiments upon affection. Let us try to do so.

Let us first of all consider the judgments of an observer O_1 upon the affective value of a set of stimuli $S_1, S_2, S_3 \cdots S_n$ to be measurements of the *ability* of these stimuli to please observer O_1. We can for short say that they are measurements of the ability A_{O_1} of stimuli $S_1, S_2, S_3 \cdots S_n$. Let us likewise consider the affective judgments of observers O_2, $O_3 \cdots O_n$ upon $S_1, S_2, S_3 \cdots S_n$ to be measurements of abilities $A_{O_2}, A_{O_3} \cdots A_{O_n}$ of these stimuli. Clearly if we establish the affective rank order of stimuli $S_1, S_2, S_3 \cdots S_n$ for observers $O_1, O_2, O_3 \cdots O_n$ and intercorrelate these rank orders we can arrange the results in the form of a table which indicates the degree of correlation between any two of the abilities $A_{O_1}, A_{O_2}, A_{O_3} \cdots A_{O_n}$ (*i.e.*, between any two of the abilities of the stimuli to please observers O_1, O_2, $O_3 \cdots O_n$). Now according to the theorem quoted above from Spearman, if the tetrad equation holds throughout such a table and only if it does, every individual measurement of every ability can be divided into two factors, a general factor 'g,' and a specific factor 's,' the former (g), although varying freely from individual to individual, remaining the same for any one individual in respect of all correlated abilities. But according to our conventions, 'ability' is 'ability to please,' and the 'individuals' are the stimuli. Thus if in a table of correlations arranged as indicated, the tetrad equation holds throughout, and only if it does, we may infer that there is a general characteristic α which, although varying freely from stimulus to stimulus, remains the same

[3] C. Spearman, The abilities of man, New York, 1927, 74–75.

for any one stimulus in respect of all correlated abilities to please the different observers—that, in brief, the stimuli involved have a general affective value for the observers involved, in spite of the divergences of the affective judgments of the individual observers.

If the considerations above be legitimate, the application of Spearman's method to affection enables one not only to ascertain the existence of a general affective factor in the case of a given class of stimuli, but also to ascertain the degree of generality of this general affective factor. This degree of generality is clearly equal to the proportion of unselected observers whose affective ranking of the stimuli when intercorrelated yields coefficients of correlation which satisfy the tetrad equation.[4] Should it be found that any group of observers chosen at random satisfies these conditions, the general affective factor may be considered to be entirely general.

But further—provided always that our theoretical considerations be correct—the application which we are proposing for Spearman's method enables us to ascertain not only the existence and generality of a general affective factor in the case of any class of stimuli, but the degree to which it is operative in the judgments of any given observer and its value in the various individual stimuli. This information can be secured very easily by applying to appropriate affective data the formulæ proposed by Spearman for computing the correlation between a specific ability and '*g*,'[5] and for computing the amount of '*g*' of a specific individual.[6]

It might be added that this line of thought suggests a ready means of distinguishing æsthetic affective value from non-æsthetic—affective value being considered æsthetic in proportion as it is general in the sense in which the term is used above. It also suggests a means of distinguishing between the æsthetically sensitive and the æsthetically dull. Given a set of æsthetic objects, *i.e.* objects to which there

[4] Provided that account be take of the influence of overlap of specific factors. *Cf.* Spearman, *op. cit.*, 150*ff* and Appendix, VII–VIII.

[5] C. Spearman, *op. cit.*, Appendix, XVI.

[6] *Ibid.*, Appendix, XVII.

have been assigned entirely general affective values in the sense in which the term is used above, the æsthetic sensitivity of an individual to these objects could be considered a function of the correlation between the affective values assigned by him to the objects and their general affective values.

Before closing this paper I wish to supplement the abstract exposition of the method which I am proposing by giving a concrete example of its application. In the course of an experiment upon affective habituation carried out at Harvard in 1925–26, I had occasion to determine the affective rank orders of 14 olfactory substances for each of 8 observers.[7] Intercorrelation of these rank orders yielded the Spearman coefficients given in the table below.[8]

TABLE I

	Ad.	Br.	Bu.	De.	Ka.	Ke.	Wa.	Ze.
Ad.		.62	.30	.40	.60	.37	−.21	.27
Br.	.62		.74	.81	.72	.69	.35	.45
Bu.	.30	.74		.66	.66	.88	.58	.38
De.	.40	.81	.66		.76	.64	.53	.66
Ka.	.60	.72	.66	.76		.73	.48	.45
Ke.	.37	.69	.88	.64	.73		.63	.54
Wa.	−.21	.35	.58	.53	.48	.63		.33
Ze.	.27	.45	.38	.66	.45	.54	.33	

The question arises: Do these correlations indicate that olfactory substances have a general affective value which determines in part the affective judgments of these eight observers? In order to answer this question we determined the 420 tetrad differences involved in the intercorrelation of 8 variables and compare their distribution around zero to the distribution which could be expected from errors of

[7] The method was that of paired comparison. The stimuli were the substances listed in Table 3.

[8] These coefficients, calculated by the method of rank differences, are treated below as though they had been calculated by the product-moment method, although it has been shown that coefficients calculated from the same data by the two methods are not necessarily the same. (*Cf.* K. Pearson, On further methods of determining correlation, Drapers' Company Research Memoirs, Biometric Series, IV., 1907, 9*ff*.) The difference involved is so slight, however, that it has not been thought necessary to take it into account in computations which are merely for illustrative purposes.

sampling for a set of tetrad differences equal in number and of like probable error. The tetrad differences t.d. are given by the formula

$$\text{t.d.} = r_{ap} \cdot r_{bq} - r_{aq} \cdot r_{bp},$$

in which a, b, p and q represent four abilities and r_{ap} represents the correlation between abilities a and p, r_{bq} the correlation between abilities b and q, etc.[9] The probable error is best given, according to Spearman, by the formula

$$\text{p.e.} = \frac{1.349}{N^{1/2}} [r^2(1 - r)^2 + (1 - R)S^2]^{1/2}$$

in which

$$R = 3r \frac{n - 4}{n - 2} - 2r^2 \frac{n - 6}{n - 2},$$

N is the number of individuals tested, n the number of abilities involved, r the mean of the correlations taken into account, and S^2 the mean squared deviation of all the correlations from their mean.[10] In Fig. 1, the observed distribution is indicated by a continuous line; the theoretical distribution to be expected from errors of sampling alone is indicated by a broken line. The two distributions are practically identical. According to Spearman, this indicates that the table satisfies the tetrad criterion and consequently if our previous theoretical considerations be correct, we may infer the likelihood that olfactory substances have a general affective value which we may call α_{oef}. I say 'the likelihood' because it is clear that both the number of olfactory substances and the number of observers were too small in this experiment to allow a high degree of probability in the conclusions. This deficiency, however, can easily be remedied by more extensive investigation.

The fact that our first question is answered positively at once raises two further questions: (1) To what extent does the general affective value α_{oef} of our olfactory substances determine the affective judgments of our observers; and

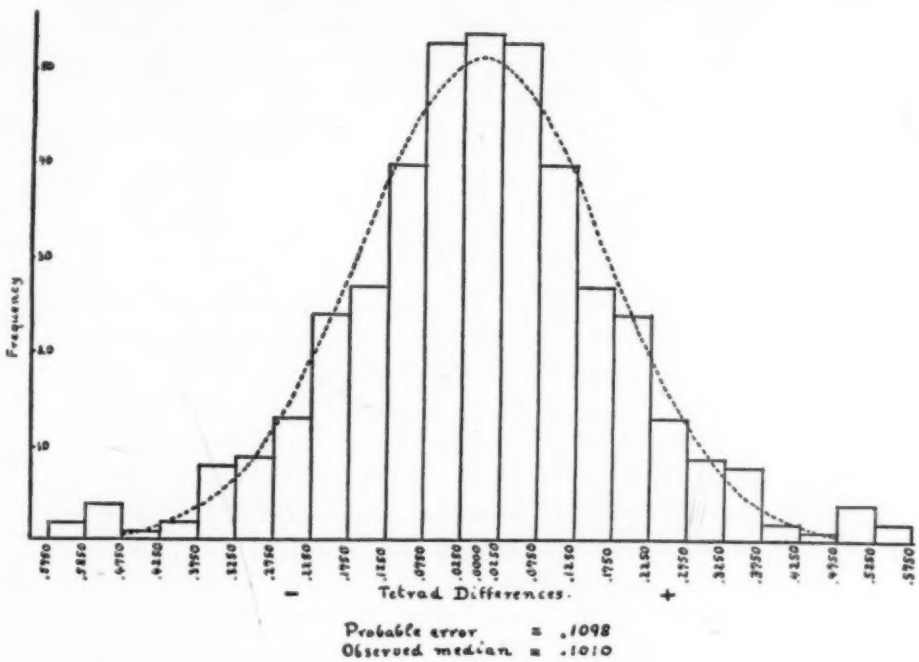
[9] C. Spearman, The abilities of man, New York, 1927, 73.

[10] *Ibid.*, Appendix, XI.

(2) What are the general affective values of our olfactory substances?

The first question may be answered as follows: According to Spearman, the correlation between the general factor 'g' and any ability a is given by the formula [11]

$$r_{ag} = (A^2 - A')^{1/2}/(T - 2A)^{1/2},$$



where "A is the sum of the correlations between a and every other test, A' is the sum of the squares of these correlations, and T is the total of all the correlations in the whole table." [12] But in our case 'g' is the general affective value of the olfactory substances and a is the ability of these substances to please a given observer. This equation consequently enables us to calculate the correlation between the general affective value of the substances and their affective value to the various individual observers, and thus to ascertain the part which the

[11] This formula holds only when all the correlations throughout a table obey the tetrad equation.

[12] C. Spearman, *op. cit.*, Appendix, XVII.

general affective value of the substances plays in the judgments of the various observers. Application of the equation yields the following results.

TABLE 2

Observers	Correlation between General Affective Value of Olfactory Substances and their Affective Value to Particular Observers
De.	.89
Ke.	.89
Ka.	.88
Br.	.87
Bu.	.83
Ze.	.58
Wa.	.48
Ad.	.41

Clearly if our theoretical considerations be correct and if one disregard the fact that our data are very limited in scope, it is possible to infer that observers De. and Ke. judge olfactory substances far more in accord with their general affective value than do observers Wa. and Ad. From this in turn it is possible, if one accept the definition of 'æsthetically sensitive' given above, to state that De. and Ke. are æsthetically more sensitive to odors than are Wa. and Ad.

The second question raised above, that concerning the general affective values of the olfactory substances involved in the experiment, may be answered as follows. According to Spearman, the amount of 'g' for any individual x tested by a battery of tests $a, b \cdots z$ is proportional to G_x as defined by the equation

$$G_x = w_a \cdot m_{ax} + w_b \cdot m_{bx} + \cdots + w_z \cdot m_{zx},$$

in which $m_{ax}, m_{bx} \cdots m_{zx}$ indicate the scores of individual x in tests $a, b \cdots z$; and $w_a, w_b \cdots w_z$ indicate optimal weights to be given to the respective scores.[13] These optimal weights in turn are best determined according to Spearman by the formula

$$w_u = r_{ug}/(1 - r^2_{ug}),$$

[13] *Ibid.*, Appendix, XVIII.

in which u is any one of tests $a, b \cdots z$.[14] But in our case 'g' is the general affective value of odors and $a, b \cdots z$ are the abilities of these odors to please different observers. These equations consequently enable us to ascertain the closest approximation possible from our data for the general affective values of the different olfactory substances used by us. Application of these equations yields the following general ranks for the 14 olfactory substances used in the experiment (the measurement of the affective values of the odors in the experiment was by the method of paired comparisons and yielded only rank orders):[15]

TABLE 3

Rank	Olfactory Substances
1 (most pleasant)	Sweet Orange
2	Extract of Carnation
3	Oil of Jasmin
4	" " Bergamot
5	" " Ylang
6	" " Cloves
7	" " Ceylon Cinnamon
8	" " White Rose (synthet.)
9	" " Neroli bitter
10	" " Cananga
11	" " Petit grain
12	" " Thyme
13	" " Rosemary
14 (least pleasant)	" " Geranium

Before closing this paper I wish to reiterate that the results given above are merely for the purpose of illustration. The data from which they are computed are clearly insufficient to give any factual standing. I also wish to point out that the general method proposed, if valid for the field of affection,

[14] *Ibid.*, Appendix, XIX.

[15] It is interesting to note that this rank order agrees very highly with that determined by the mere summing of the ranks assigned to the odors by the different observers. The correlation between the two rank orders is .98 by the Spearman formula. It is clear that the summing of ranks, whether weighted or not, introduces a large source of error when these ranks are considered to represent, as in this case, serial position with respect to an intensive continuum. The same source of error is present, however, in the use of Spearman's rank correlation formula, and this formula has nevertheless been of great value to psychology.

must likewise be valid for all other fields involving judgments of value, whatever their nature. In consequence the application of this method to æsthetics in no wise requires the acceptance of a hedonistic point of view with respect to that science.

[MS. received April 12, 1929]

THE CONDITIONED REFLEX AND THE SIGN FUNCTION IN LEARNING

BY KATHERINE ADAMS WILLIAMS

Radcliffe College

No single psychological principle has had a greater vogue in recent years than that of the conditioned reflex. While it has not received the unqualified approval from all psychologists which it has received from a few, and while it has not been immune from criticism, yet it has found a place in almost every textbook written in the last fifteen years and flourishes in some quarters to the virtual exclusion of most other fundamental psychological concepts. The present paper would submit the conditioned reflex to a closer examination, outline some neglected considerations, and on the grounds indicated, challenge the doctrine's adequacy as either a descriptive concept or an explanatory principle in the more important cases to which it has been applied. Before this can be done, it will be necessary to review a little of the history of the concept, its original formulation, and the subsequent uses to which it has been put.

Attention was first called by Pavlov to the phenomenon which he designated the 'conditioned reflex.' He found that a stimulus which originally did not bring a given reflex into action, could be made to do so by repeatedly presenting with it another stimulus which was so effective from the first. The originally inactive stimulus he called the conditioned stimulus, and the originally active one, the unconditioned stimulus. The reflex called forth by the conditioned stimulus was qualitatively identical with the reflex to the unconditioned stimulus. For instance, in the case of the salivary reflex, if the unconditioned stimulus was meat, the conditioned stimulus produced a flow of saliva such as meat produced, whereas if the unconditioned stimulus was bread-crumbs, the conditioned stimulus produced saliva of the sort which

bread-crumbs produced. The quantity of saliva secreted might be less for the conditioned stimulus, but the quality was the same. In general Pavlov insists that the conditioned reflex has exactly the same mechanical and determinate character which, in his view, is typical of the original or unconditioned reflex.[1] Various features of the conditioned reflex which give it great flexibility have been experimentally analyzed, but these, to Pavlov's mind, only serve to show the fineness of the mechanism, not to make it less mechanical.

Apparently any stimulus whatever may be made a conditioned stimulus. The Pavlov school has used dogs almost exclusively, and has devoted most of its attention to the salivary reflex of the alimentary type, the unconditioned stimulus being food; or to the salivary reflex of the defensive or protective type, the unconditioned stimulus being acid introduced into the mouth. Some indifferent stimulus such as a bell or a light is repeatedly presented just before the dog is given food. After a number of such trials, the bell or the light alone is sufficient to call forth a copious flow of saliva, and the reflex is said to be conditioned. Stimuli in almost every sense mode have been used as conditioned stimuli, the modes most commonly employed being the auditory, the visual, and the tactile. Furthermore, so far as is known, any response may be conditioned. As has been said, Pavlov has used the salivary reflex. Cason, in his experiment on the pupillary reflex, succeeded in obtaining evidence that smooth muscle reflexes can be conditioned.[2] Bechterew,[3] and later, Watson,[4] conditioned a striped muscle reflex, namely, a flexion or withdrawal reflex. Since muscular and glandular, and voluntary and involuntary, reactions have been conditioned, it seems likely that any reflex can be.

On account of the supposed mechanical nature of the

[1] I. Pavlov, Conditioned reflexes, translated by G. V. Anrep, Oxford University Press, 1927, p. 24.

[2] H. Cason, The conditioned pupillary reaction, J. Exper. Psychol., 1922, 5, 108–146.

[3] V. v. Bechterew, La psychologie objective, translated by N. Kostyleff, Paris, Alcan, 1913, p. 253.

[4] J. B. Watson, The place of the conditioned reflex in psychology, Psychol. Rev., 1916, 23, 89–116.

conditioned reflex, and its widespread occurrence, it has been seized upon by the 'tough-minded' as affording the central principle of psychology. Innate reflexes, it is held, account for the innate equipment of man and animals, and the conditioned reflexes account for modification.

The cue for this thoroughgoing acceptance of the conditioned reflex principle was given by Pavlov and Bechterew, the two great pioneer investigators, the one of secretory reflexes and the other of motor reflexes. Pavlov believes that all behavior is reducible to reflexes and combinations of reflexes. He examines the claims of 'instincts,' finds all of them wanting, and concludes, "The aggregate of reflexes constitutes the foundation of the nervous activities both of men and of animals." [5] The result is that we hear of the investigatory reflex, which receives the following description: "It is this reflex which brings about the immediate response in man and animals to the slightest changes in the world around them, so they immediately orientate their proper receptor organ in accordance with the perceptible quality in the agent bringing about the change, making full investigation of it." [6] The alimentary reflex and the "mild defense reflex to rejectable substances" have two components. "Firstly, the animal exhibits a reflex activity directed towards getting hold of the food and eating it, or, in the case of rejectable substances, toward getting rid of them out of the mouth; and secondly, in both cases, an immediate secretion of saliva occurs. . . ." [7] On this foundation the learned behavior of the animal is built up of conditioned reflexes. "We recognize them in ourselves and others under such names as 'education,' 'habits,' and 'training'; and all these are nothing more than the results of an establishment of new nervous connections during post-natal existence of the organism." [8]

Bechterew is no less sweeping. Indeed, he is perhaps more so, for after stating that the 'réflexes associées' are as

[5] I. Pavlov, *op. cit.*, p. 11.
[6] I. Pavlov, *op. cit.*, p. 12.
[7] I. Pavlov, *op. cit.*, p.17.
[8] I. Pavlov, *op. cit.*, p. 26.

mechanical and simple as the simple reflexes, he says, "Toute notre pensée se reduit à des réflexes de ce genre." [9] More particularly such 'mentalistic' phenomena as association of ideas, judgment, comparison, and inference, are reducible to conditioned reflexes.

John B. Watson in 'Behaviorism' accords the concept a similarly important place. He says, "The relationship, theoretically, between the simplest cases of the conditioned responses and the more complicated, integrated, spaced and timed habit responses we are considering tonight, seems to me to be quite simple. It is the relationship of part to whole—that is, the conditioned reflex is the unit out of which the whole habit is formed." [10] Not only this, but the integrating, the patterning, the timing, seem themselves to be processes of conditioning. It is apparently correct to say that the linking of a simple reaction to a simple stimulus, the linking of simple reactions to each other to form a complex response, and the linking of a complex response, such as a habit or an emotion, to any stimulus, are all cases of conditioning. In a word, conditioning becomes synonymous with learning.

The concept has, in addition, been more specifically applied as an explanatory principle. A comparatively modest use of the conditioned reflex as a principle of learning is made by Allport. He begins with certain innate prepotent reflexes. "The prepotent reflexes are subject to synaptic changes in their central portions. The effects of such changes are (1) to extend the range and complexity of the stimuli capable of exciting the response, and (2) to refine and specialize the response itself. The first effect, which may be called an afferent modification, is brought about by the principle of the conditioned response; the second, resulting in an efferent modification, is due to the selection and fixation of successful random movements in the processes of habit formation and thought." [11] The prepotent reflexes are at-

[9] V. v. Bechterew, *op. cit.*, p. 267.

[10] J. B. Watson, Behaviorism, New York, W. Norton and Co., 1924, p. 68.

[11] F. H. Allport, Social psychology, New York, Houghton, Mifflin Co., 1924, p. 56.

tached to simple stimuli to which an infant is adapted to respond. It is due to the afferent modifications that the responses of the adult can be called forth by an enormous number of persons, things, and situations.

Smith and Guthrie go farther, citing the conditioned response as furnishing the principle of selection of learned responses, the most difficult point in the explanation of learning.[12] Frank [13] and Wilson [14] have followed them. They take the case of a cat in a puzzle box. The question is, how does the cat after repeated trials come to eliminate all the useless reactions and to retain only the successful one of pulling the latch which leads to release? The reply of these authors is that this comes about simply through the conditioning of various reflexes. At first the cat is attracted by the sight of the bars, say, and bites at these. But pussy has an innate aversion to the sensation of biting unyielding objects, and so withdraws from the bars. The result of numerous repetitions of this bit of behavior is that the sight of the bars becomes a conditioned stimulus for the withdrawing. The only thing which does not take on such negatively conditioned properties is the latch. The cat, therefore, when placed in the box has nothing left to do but to pull the latch.

Such applications of the concept have not been spared attack. These attacks have been made along two main lines corresponding to the two components of the concept 'conditioned reflex.' Some psychologists, whether or not they have accepted the concept of conditioning, have criticized the wide use of the concept of the reflex, while others, whether or not they have accepted such wide use of the reflex, have attacked the concept of conditioning.

The criticism directed primarily against the concept of the reflex has been of two sorts. In the first place, it has urged

[12] S. Smith, and E. R. Guthrie, General psychology in terms of behavior, New York, Appleton, 1921, pp. 75-133.

[13] L. K. Frank, Suggestion for a theory of learning, PSYCHOL. REV., 1923, 30, 145–148.

[14] W. R. Wilson, Selection in trial and error learning, PSYCHOL. REV., 1924, 31, 150–160.

that the implications of the concept are contradicted by the facts. For instance, Lashley's 'Studies of Cerebral Function in Learning' I–VII,[15] have done much to show that reflex paths cannot be so perfectly determinate as the reflex theories demand.

In the second place, the adequacy of the concept of the reflex as a description of what actually happens has been contested. This protest may be found in the writings of McDougall, in those of Tolman, and in those of the Gestalt school. If, for example, we examine Pavlov's description of the alimentary reflex quoted above, we do not find anything which fulfills his definition of a reflex. Apparently the only character which can identify such reflexes is that they serve to bring about a certain readjustment between the environment and the organism. This describes a purposive instinct and not a mechanical reflex.

The criticism which attacks the concept of conditioning is of greatest interest here. This, like that directed against the concept of the reflex, may take two main lines. Most attention has been paid to the question whether certain special implications of 'conditioning' as a principle to explain learning are in agreement with the facts. If, as the theory demands, the learning of the maze is really the learning of a series of conditioned reflexes, errors ought to be eliminated in a backward direction from the food-box. In other words, the alley leading to the food-box becomes a conditioned stimulus for approach; the alley preceding that, becomes a secondarily conditioned stimulus for approach to the last alley, and so on back to the first alley. Helen Hubbert,[16] Stella Vincent,[17] Hubbert and Lashley,[18] and Carr [19] analyze

[15] For bibliography of these, see E. C. Tolman, Habit formation and higher mental processes in animals, PSYCHOL. BULL., 1927, 24, 1–35.

[16] H. Hubbert, Elimination of errors in the maze, *J. Animal Behav.*, 1915, 5, 66–72.

[17] S. Vincent, The white rat and the maze problem, IV: The number and distribution of errors, a comparative study, *J. Animal Behav.*, 1915, 6, 367–374.

[18] H. Hubbert and K. S. Lashley, Retroactive association and elimination of errors in the maze, *J. Animal Behav.*, 1917, 7, 130–138.

[19] H. Carr, The distribution and elimination of errors in the maze, *J. Animal Behav.*, 1917, 7, 145–159.

data actually obtained in mazes to discover whether this happens. The outcome of the discussion is rather indecisive. No marked backward elimination appears. It is possible, however, that there is some such effect which is obscured by other factors.

The second line of criticism, which would question the adequacy of the concept of conditioning to describe what actually occurs, has been comparatively neglected. Yet upon this adequacy the whole case for the conditioned reflex must rest, in the present state of our knowledge, for as we have seen, the appeal to necessary implications yields no conclusive results. We shall examine a case of learning in some detail with a view to seeing whether the facts really fit the conditioned reflex scheme.

A rat, let us say, has learned to choose, in a black-white discrimination box, the white compartment, in which it finds and eats food, rather than the black compartment, which is empty. The most uncritical account of what happens in the discrimination box assumes that food is the unconditioned stimulus, *S* (Diagram 1), to the response of approach and

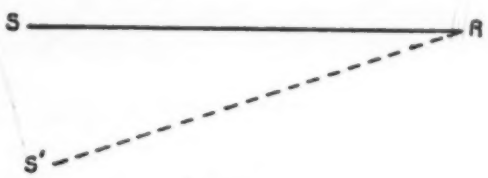


DIAGRAM 1

eating, *R*. When food is repeatedly presented to the animal in the white compartment of the box, this white compartment becomes a conditioned stimulus, *S'*, to the response made to food. This analysis, however, is obviously too crude. 'Food' is not a single stimulus but a complex of stimuli, and some of these stimuli affect the animal at one time in one way, some at another time in another way.

A second analysis might say that the appearance and odor of food are the unconditioned stimuli, S_1 (Diagram 2), for the approach response, R_1. Approach brings new stimuli to bear, such as tactual and gustatory stimuli, S_2, which in

turn call forth the response of eating, R_2. The white discrimination compartment, under our conditions, always accompanies the appearance and odor of food, and so becomes a conditioned stimulus, S', which may be substituted for

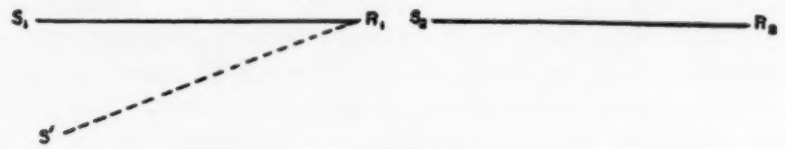


DIAGRAM 2

the appearance and odor of food. Now, with rats, no work has been done on this point, but experimental evidence shows that in the case of dogs, at least, the sight and smell of food do not arouse approach responses unless eating has followed in the previous experience of the animal. In other words, the visual and olfactory stimuli from the food are not unconditioned stimuli but are themselves conditioned stimuli. Puppies fed entirely on milk since birth failed to respond to the sight and odor of meat until they had eaten it several times.[20]

A third attempt at analysis may say, then, that the white box, S'' (Diagram 3), and the appearance and odor of food,

DIAGRAM 3

S', are all conditioned stimuli to the response of eating, R, whose original unconditioned stimulus is the tactual or gustatory stimulus, S. But to this analysis, too, there is a

[20] C. L. Evans, Recent advances in physiology, 3rd. ed., Phila., Blakiston's, 1928, pp. 261-262.

decisive objection. The rat does not respond to the white box by eating; he can't, for there is nothing which he can eat. Conversely, he never responds to gustatory sensations by approach; he can't, for he is already there.

A scheme which would avoid this difficulty may be derived from Allport's discussion.[21] This description would be substantially as follows: The rat, whatever the environmental stimuli, responds with exploration, R_1 (Diagram 4), when he

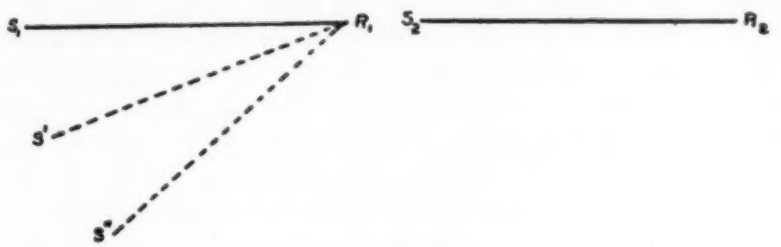


DIAGRAM 4

is experiencing the internal stimulus of hunger, S_1. In the course of these explorations, he happens upon food, S_2, and eats it, R_2. The sight of food becomes a substitute for hunger; that is, a conditioned stimulus, S'. Now if we arrange things so that when he is released for his explorations, he always happens upon a white box, the white box will also become a conditioned stimulus, S'', to exploratory approach.

But this account, too is unsatisfactory, for consider that the connection between the visual stimulus, S' or S'', and the response, R_1, is not made unless the process is consummated and rewarded by the segment, $S_2 - R_2$; that is, the rat learns to approach the white compartment not when he has merely stumbled upon it when he was hungry, but only when he has repeatedly obtained food in it.[22]

Actually the white compartment is the only area of the situation where he is anything but hungry. If the approach to the white compartment is a case of conditioning, then, it is also something more. The conditioned reflex formula

[21] F. H. Allport, *op. cit.*, pp. 61–63.

[22] K. A. Williams, The reward value of a conditioned stimulus, *University of California Publications in Psychol.*, 1929, 4, 31–55.

takes account only of the unconditioned stimulus, the conditioned stimulus, and the response. There is no room in it for S_2 and R_2, the consummatory phases, without which the conditioning, in this scheme, does not occur.

Now there is an inclusive objection to all the foregoing schemes. We have all along had tacitly to assume that there was a definite response, which might be called 'approach,' which was at first made to some part of the situation and later became attached to another part. But, as a matter of fact, the most careful examination fails to discover in the early stages of learning, among the unconditioned responses, anything whatever which remotely resembles the swift sure scramble into the white compartment, the response 'conditioned.'

We saw earlier that the conditioned reflex concept demands that the same response should be made to the conditioned stimulus that was made to the unconditioned stimulus. To be sure, the response may be quantitatively weaker. In the case of motor responses, for instance, the movement may be reduced in vigor, but if the concept of 'the conditioned reflex' is to have any distinctive and precise meaning, we must require that the movements of the conditioned response involve the same muscles as the unconditioned response. If the response in question is not so identified, it can be identified only by the end reached, and conditioning becomes a case of purposive learning. Now, our rats, certainly, do not make the same movements in the sense demanded by the conditioned reflex interpretation, nor even a weak copy of them. They make a new set of movements, and instead of a weaker set, they make a stronger set. In this case, then, there is no response made to any unconditioned stimulus whatsoever with which the conditioned response can be considered identical. The conditioned response scheme is, therefore, inapplicable.

The notion of the conditioned reflex, alleged as an 'explanatory principle,' fails to be adequate even as a general description. The remainder of this paper will develop an account which has at least the scientific merit of accommo-

dating the facts and of leaving the way open to scientific explanation of them.

To begin with, it is necessary to go beyond the specious simplicity of our previous analyses, making several important distinctions. In the first place, we must recognize that there are responses which do not require support from their stimuli, and that, on the other hand, there are responses which do require support from their stimuli. For instance, salivation takes place when food is placed in the mouth, but the food supplies no indispensable support for the process; that is, salivation can go on without anything at all in the mouth. Chewing likewise takes place when food is placed in the mouth, but unlike salivation, it does not go on without food; it requires support. Those activities which do something with the stimulus which calls them forth cannot proceed unless certain supporting qualities are offered by the stimulus.[23] It seems obvious enough that it is only the first class of activities for which stimuli can be substituted unrestrictedly. Only stimuli which have these supporting qualities can possibly serve as conditioned stimuli to supported responses.

The second distinction important for our purposes is that which has already been suggested by our discussion; namely, between responses which will never be made, or which upon repeated trials will not continue to be made, unless certain other responses follow them, and responses which will continue to be made unless highly unfavorable consequences follow upon them. We may call these types of response dependent and independent, respectively. It will readily be seen that we may similarly distinguish between responses without which the dependent responses will not occur and responses which are not indispensable in this way to the occurrence of preparatory responses. These types of response may be called requisite and non-requisite, respectively. On the whole it is likely that most dependent responses are non-requisite and that most independent responses are requisite.

[23] Prof. Tolman has distinguished between discriminabilities and manipulabilities. (E. C. Tolman, Behavioristic theory of ideas. PSYCHOL. REV., 1926, 33, 352–369.) In his terms a supported response is one whose stimulus, 'discriminability,' is also a manipulability.

But it is also possible that there should be requisite dependent responses; that is, responses the repeated occurrence of which depends upon their consequences but which are yet necessary to the occurrence of certain responses preceding them. There may also be responses of a non-requisite but independent sort; that is, responses which occur whatever the consequences (unless, of course, these are injurious) but which are not requisite for the occurrence of any dependent responses. It seems to me to be suitable to call all dependent responses preparatory, but to restrict the term consummatory to those responses which are not only independent but also requisite, and so may be regarded as the ends of chains of behavior. Independent non-requisite responses, if there be any, may be distinguished from consummatory responses by the name, isolated responses.

Whether a specific response is non-requisite or requisite preparatory, consummatory, or isolated, must be decided by observation. In the example which we have used previously in this discussion, that of the rat in the discrimination box, we have already seen that the approach to the white compartment of the discrimination box is a non-requisite preparatory response. It does not arise, in its final swift sure form, at least, and does not persist, unless there are specifiable consequences (in this case, food). The process of eating, or some part of it, is a consummatory response. It does appear and persist unless some very untoward result follows (such as an electric shock). I must turn to other situations to find unambiguous examples of the requisite preparatory and the isolated response. Suppose we discover experimentally that a dog which is prevented from swallowing, eventually ceases to salivate when food is placed in its mouth, even when the food is inserted by some other means into its stomach. This would show that swallowing is a requisite response. Without it the preparatory response of salivation will not continue to occur. But suppose that we have also shown that swallowing will not continue to occur if our dog has a gastric fistula which prevents the food from reaching the stomach. This shows that swallowing is itself a dependent or preparatory

response. It is then a requisite preparatory response. We must look still farther for an instance of an isolated response. In all probability the knee-jerk is an independent response in our sense—that is, it requires no ulterior confirmation for its continuance. Suppose, further, upon examination, we find that we can use it neither as a reward nor as a punishment; that is, that our subject will perform no action to obtain or avoid it. It is, then, both independent and non-requisite, an isolated response.

Now if there are any responses to which the attachment of new stimuli can be adequately explained by the conditioned reflex principle, the most promising place to look for them is among these isolated responses. The next most promising group of responses in which to demonstrate the adequacy of the conditioned reflex principle is that of the consummatory responses. Here, however, we are extremely likely to find that we are dealing with supported responses, so that the range of stimuli which can be substituted is comparatively narrow. In no case of the preparatory response can the conditioned reflex principle be entirely adequate, and in some cases it has no applicability at all. The conditioned reflex is helpless to account for the rise and perfection of a new response. It can deal only with the substitution of stimuli to a response already established. Furthermore, such a scheme, which takes account of the unconditioned stimulus, the response, and the conditioned stimulus, and these only, cannot comprehend cases in which one stimulus to a *preparatory* response is substituted for another, for, as we have seen, this substitution never occurs unless it is confirmed by a subsequent consummatory response. If the conditioned reflex principle plays a rôle at all here, it is certainly not the chief one.

It may assist in the handling of the problems raised by these considerations to designate the stimuli to preparatory responses as *signs*. Their character consists behaviorally in the fact that their efficacy as stimuli depends on their function of preparing the organism for something that is to come. They initiate approach and avoidance responses, not in their

own right, not even necessarily to or from themselves, but only because they have in the past been followed by certain other stimuli which are desired or feared. A second such sign cannot be substituted for a first unless the relation of the second to the consummatory situation has proved in the experience of the animal to be the same as that of the first. It is not mere temporal coincidence but identity of sign function which determines substitution of stimuli. It must now be plain that many, if not most, of the cases of learning for the explanation of which the conditioned reflex has been the sole principle called upon, cannot be understood until they are seen rather to be preparatory responses and their stimuli are seen to be effective only because of their sign character.

Pavlov himself describes conditioned reflexes as 'signal reflexes' but his whole treatment shows that by 'signal' he does not intend to convey anything not contained in the mechanical conception of the conditioned reflex. Prof. Tolman has used the term 'symbol.' [24] Ordinarily, however, a 'symbol' is a substitute for the real thing; we respond to symbols of things as we should to the things. Thus a conditioned stimulus may be regarded as a 'symbol' of the original stimulus. But the relation of a preparatory stimulus to the consummatory stimulus is best indicated by the word 'sign,' in that the meaning of a sign is not primarily the thing of which it is a sign, but some fact about this thing in relation to the sign-situation. (In our instance, the fact signified is that this thing can be obtained—or avoided—by such-and-such a preparatory response.)

The question whether there are signs which are innately effective is an experimental one. If we find that an animal in the absence of previous experience reacts to a certain stimulus but continues to do so only in case certain results follow, we conclude in favor of innately given signs. This would be a case, of course, of what has been called negative adaptation.

[24] E. C. Tolman, Habit formation and higher mental processes in animals, PSYCHOL. BULL., 1927, 24, p. 24.

It is now time to notice the objections which may be raised to the analysis here attempted. In the first place, it will be argued that the conditioned reflex is a simple, mechanical, well-established concept, and that therefore it should be given first rights; the more cases which can be brought under its egis, the farther along we shall be. My reply to this is that the claim of the conditioned reflex concept rests on its adequacy as a descriptive concept. Even supposing that it has the merits which its proponents claim, we ought unhesitatingly to substitute a less simple and mechanical and well-established concept if the latter more accurately describes the situation to which it is to be applied. This I have tried to show that the concept of signs does. But in our present knowledge of conditioned reflex, I must regard the claim of mechanical simplicity as specious, for it begins to seem that the more thoroughgoing the work on reflex paths is, the more complicated and less mechanical must be the theories about them. At least we are far from a satisfactory neurological theory. We need have no scruple, therefore, against the concept's replacement.

Two more special criticisms, oddly enough, take opposite directions. In the first place, it may be argued that the present analysis is too atomistic. We cannot properly speak of the response of approach to food, of the response of taking food into the mouth, of the response of chewing and swallowing, and so on. There is only one response and that is the response of seeking-and-finding-and-eating food. In the situation of the rat in the discrimination box, the conditioned stimulus does actually take the place of an originally active stimulus and does initiate the same identical response of seeking-and-finding-and-eating food. The only difference is that when environmental supports are absent, the response ends in mid-air, so to speak, and so is incomplete. My reply to this view is that it fails by giving up the attempt at precise analysis. It seems to me justifiable to take as *a response* whatever in the behavior of an animal becomes altered when any part of the exciting situation is altered, while *a stimulus* is any part of the exciting situation, the alteration of which

can produce a change in the response. Now for some purposes it will be justifiable to take the complex situation describable as a hungry animal in a discrimination box containing food, as the stimulus, and the complex response describable as exploration and the finding and eating of food with the resultant quiescence, as the response. However, since the various parts of the situation are capable of producing changes of various sorts in the animal's behavior, it also seems to me to be justifiable and necessary to take smaller units and to ask concerning the relation of these to one another.

The other special criticism to which I have referred is that our analysis is not minute enough. It is urged that if we only took still smaller units we should be able to recognize the conditioned reflex in all its purity. I readily, of course, admit that it is possible and perhaps desirable to look to the smaller units, but I deny that by doing so we should be able to reduce everything to conditioned reflexes. If the single response of approach to the white compartment does not occur unless it is followed by eating, we certainly can have no hope that the individual steps (the smaller units) by which the approach is made will none-the-less follow upon their individual stimuli. Neither are we helped by minuter analysis in finding an original response identical with the final learned response, for after all, a series of parts repeated in an identical order would give us an identical whole.

Finally, it may well be asked what all this accomplishes. To my mind it accomplishes a modest but necessary thing. It contributes to the analysis of the conditioned reflex concept, showing its inadequacy as it is widely used to describe learning, and, therefore, to explain it. It substitutes another concept which, although it cannot offer itself as in any way an explanation of learning, may yet hope by its congeniality with such purposive theories as that of Prof. Tolman to further such explanation.

Summary

1. The apparently simple and mechanical character of the conditioned reflex so encourages the assumption of its useful-

ness that many psychologists have advanced it as an explanation of the learning process. This acceptance the writer believes to be a mistake.

2. The conditioned reflex, since it has no assured neurological explanation which would give it preference, must rest its case on its adequacy as a concept to describe behavior, and here, when we appeal to actual cases, we find its application very limited.

(*a*) Since it takes account only of substitution of stimuli, it cannot comprehend cases in which a new response arises and is perfected. Yet many of the responses in question cannot be regarded as other than new.

(*b*) Since it takes into its scheme only one response, the original stimulus to this, and the substituted stimulus, it cannot account for the contingency of the establishment and continuance of preparatory responses upon subsequent consummatory responses.

(*c*) For the same reason, the concept is inadequate for all those cases of substitution of one stimulus to a response for another which are contingent upon subsequent consummatory responses.

3. Analysis discloses that the essential feature of learning is not the substitution of stimuli by virtue of coincidence of the original stimulus and the substituted stimulus, but the contingency of the learned response upon a subsequent response. This relation of preparatory stimulus to consummatory stimulus can be identified and described as the 'sign' function, and its recognition will further a scheme of real explanation of the learning process.

[MS. received May 1, 1929]

A FUNCTIONAL INTERPRETATION OF THE CONDITIONED REFLEX

BY CLARK L. HULL

Yale University

Introduction

The experimental evidence now available shows quite clearly that the conditioned reflex is a two-phase phenomenon. One phase is obviously primary and the other is definitely secondary. Viewed physiologically, the primary phase is positive or excitatory in its nature; the secondary phase is negative or inhibitory. Functionally regarded, the primary phase appears to be a tentative trial, or first-approximation aspect of an adaptative process, while the secondary phase is the selective, corrective, or precision-insuring aspect. These two phases of the conditioned reflex, operating jointly, thus stand revealed as an automatic trial-and-error mechanism which mediates, blindly but beautifully, the adjustment of the organism to a complex environment.

The primary or excitatory phase of the conditioned reflex is the one which is best known and which has been employed most extensively as an explanatory principle. The knowledge of certain aspects of it, indeed, is as old as associationism itself. A much more comprehensive view of the process has recently been exploited to considerable advantage by H. L. Hollingworth under the name of *redintegration.* As applied to the conditioned reflex, this term represents the fact that all elements of a stimulus complex playing upon the sensorium of an organism at or near the time that a response is evoked, tend themselves independently and indiscriminately to acquire the capacity to evoke substantially the same response. For our present purposes the indiscriminateness of the tendency is particularly to be noted.

But the redintegrative aspect is only one of at least four which are discernible in the primary phase of the conditioned

reflex. A second significant tendency is an almost total lack of responsiveness to the patterning of the stimulus complex. A third is a remarkable lack of specificity of the conditioned reactions as regards the conditioned stimuli which may evoke them; the reflexologists call this *irradiation*. A fourth characteristic of great significance is the curious tendency, where the conditioned stimulus precedes the unconditioned one in the conditioning process, for the reaction to be attracted forward toward the former. Under certain circumstances the reaction (after a number of reenforcements) may begin a considerable interval before the delivery of the unconditioned stimulus. In the case of certain defense reactions this may even result in the organism not receiving the nocuous unconditioned stimulus at all.

The secondary or inhibitory phase of the conditioned reflex appears to be less widely appreciated. For the most part this phase is not open to ordinary observation, only becoming manifest as the result of ingenious experimental procedures. Corresponding to the four aspects of the excitatory phase, each to each, we find here four parallel inhibitory aspects. They are: (1) inhibition from experimental extinction, (2) conditioned inhibition, (3) differential inhibition, and (4) inhibition of delay. In this connection it is to be noted that a given inhibitory tendency can only be developed on the basis of a corresponding excitatory tendency which must previously have been established.

We may now proceed to the consideration of the biological function performed by the several phases of the conditioned reflex process.

Redintegration and Experimental Extinction

Of what biological utility is the redintegrative tendency? It clearly results in the multiplication of the stimulus complexes which are capable of evoking particular reactions. With certain limitations, these conditioned stimulus complexes become equivalent to, *i.e.*, substitutable for, the corresponding native or unconditioned stimuli. But just how does this substitution tendency result in augmenting the

survival chances of the organism? It is quite clear, for example, that for any and every stimulus complex to have the capacity to evoke any and every response would not be good biological economy. Such an arrangement could lead to nothing but a wild and unadaptive chaos of behavior. No doubt many psychologists and biologists with a vitalistic leaning will urge that, if the process be really blind and automatic as assumed, we should expect exactly such a chaos. The problem deserves serious consideration.

The solution of the problem is seen perhaps most readily in the conditioning of defense reactions. The unconditioned stimuli for such reactions are ordinarily genuine injuries. With such unconditioned stimuli the organism will rarely or never make an unnecessary defense reaction for the reason that a defense will always be needed. Such certainty could hardly be attained with any other type of stimulus. This is a characteristic example of biological conservatism. The trouble with this particular type of arrangement is that, in order for the defense reaction to take place, the organism must always receive an injury. This is bad biological economy. Clearly a corrective accessory mechanism is needed. This exists in the substitution-of-stimulus tendency characteristic of redintegration.

Now the nature of nocuous stimuli practically limits them to such as involve actual contact with the organism before being effective. But if, as will usually be the case, the nocuous stimulus is of such a nature as also to stimulate a distance receptor like the eye, this latter stimulus is likely to get conditioned to the defense reaction. Here we have a means whereby effective defense behavior may be evoked without always being preceded by an injury. The retinal image of the threatening object when at a moderate distance will be sufficiently like that which is received when it is close enough to deliver the injurious stimulus, to evoke the defense reaction (withdrawal, flight) early enough for the organism to escape the injury altogether. Indeed it may very well be that the frequency among primitive conditioned reflexes of the substitution of distance receptors for contact receptors is due to this combination of circumstances.

Granting the tremendous biological advantage of occasionally being able to substitute certain stimulus complexes for certain others we still are pursued by the threat of a behavior chaos. There remains, in short, the difficulty presented by the indiscriminateness of the redintegrative tendency. Quite irrelevant stimulus elements will almost certainly find their way into every stimulus complex. By the principle of redintegration alone these irrelevant ones must get conditioned exactly as do the relevant. Why does not this produce the blind chaos of behavior previously suggested?

The answer is found in the corrective principle of experimental extinction. Stimulus elements which are not biologically relevant will not accompany a given unconditioned stimulus with any regularity, whereas the truly significant elements must do so. The latter, of course, will develop ordinary conditioned reflexes. The former, also, will tend to do so during their first accidental reinforcements or occasional short unbroken sequences of reinforcements. In so far as this accidental reinforcement takes place there may be realized a genuine unadaptiveness of behavior. Presumably this mechanism is responsible for a certain amount of human and other animal error.

Fortunately complete functional conditioning usually does not take place until after repeated combined stimulations. Except for very unusual runs of chance coincidences of stimuli, the irrelevant stimulus would appear one or more times *unaccompanied* by the unconditioned stimulus before the accidentally initiated redintegrative tendency should have risen above the functioning threshold. Such failures of reinforcement at once produce a tendency to experimental extinction. In this connection it must be remembered that experimental extinction is not a mere passive failure to strengthen an excitatory tendency according to the so-called 'law of use.' Instead it is a very potent tendency to repress existent excitatory tendencies, particularly the one from which it has taken its origin.[1] Since chance alone will ordinarily present the irrelevant stimulus without reinforcement much

[1] See I. P. Pavlov's Conditioned reflexes, pp. 54*ff* (Oxford University Press, 1927).

more frequently than with it, the resulting inhibitory tendency will very soon become much more potent than the positive redintegrative tendency. Even if by some chance the false conditioned tendency should have gotten above the reaction threshold, the combination of circumstances just referred to would very soon convert it into a permanently inhibited and impotent state.

Patterned Stimuli and Conditioned Inhibition

After observing the utter indiscriminateness of the primary phase of the conditioned reflex as to the components of the stimulus complex which it tends to endow with action-evoking powers, we should not be greatly surprised to find a similar obtuseness as regards sensitivity to the particular combination or patterning of such complexes. Extensive experiments show, as a matter of fact, that the primary conditioning tendency leaves the components of the conditioned stimulus in an essentially unorganized state as regards the evocation of response. It is true that, if only a part of the original stimulus complex be presented, the intensity and promptness of the response will be reduced. This, however, is an addition-subtraction type of reaction rather than a sensitivity to organization or pattern. Barring accidental variability in the potency of the several components of the conditioned stimulus, this reduction in the magnitude of the response closely parallels the reduction in the number of the conditioned stimulus elements. With the same reservation, it may be said that one combination of stimulus elements from an original conditioned stimulus complex, will evoke the same response (both qualitative and quantitative) as any other combination having the same number of elements. Similarly, if two distinct stimuli which have been independently conditioned to a given response be presented together, the intensity of the resulting response is likely to approach closely the arithmetical sum of the responses to the two stimuli if presented separately. It is accordingly clear that, except for characteristic differences in potency, the individual components of a primarily conditioned stimulus

complex are completely interchangeable and appear to have little or no functional individuality. Under such circumstances there is naturally no differential sensitivity to any particular combination or pattern of stimulus components.

Now it is evident to ordinary observation that the simple addition-subtraction relationship obtaining among the components of a conditioned stimulus in the primary phase of the conditioning process, is a fairly adequate first approximation for many life situations. Indeed, if the vertebrate organism were to be dependent upon but a *single* stimulus mechanism, it is doubtful whether any other conceivable one would be more conducive to successful environmental adjustment and survival. In the long run, where fewer signs of danger appear, the less danger there is likely to be. Similarly, where two signs of danger appear, both of which independently are tolerably reliable, the organism is pragmatically justified by the law of chance alone in making unusually prompt and vigorous defense reactions. The same may be assumed to hold for positive reactions such as those involved in food getting.

Even so, innumerable life situations arise where the simple addition or subtraction of the potencies of the several components of a stimulus complex is not adequate. In many situations a particular combination or pattern of stimulus components (either simultaneous or temporally extended) is the very essence of the stimulus. To change a single minute component of certain stimuli will completely change the nature of the appropriate response. A telegram is an example of such a patterned stimulus complex. If a single letter in it be changed, the reaction of the receiver may be made either one of joy or of despair.

Numerous experimental examples of differential sensitivity to the patterning of stimuli are found in the conditioned reflex literature. We reproduce from Pavlov one involving a temporal pattern:

> The following is an experiment by Dr. Ivanov Smolensky. The positive conditioned alimentary stimulus was made up of a hissing sound (H), a high tone (hT), a low tone (lT), and the sound of a

buzzer (B), applied in that order, namely H–hT–lT–B. The inhibitory stimulus was made up with the order of the two middle components reversed, namely H–lT–hT–B.

Time	Conditioned Stimulus	Secretion of Saliva in Drops During 30 Seconds	Remarks
3:10 P.M.	H–hT–lT–B	4	Reinforced
3:17 P.M.	H–lT–hT–B	0	Not reinforced
3:27 P.M.	H–hT–lT–B	3	Reinforced
3:32 P.M.	H–hT–lT–B	4	"
3:38 P.M.	H–lT–hT–B	0	Not reinforced
3:46 P.M.	H–hT–lT–B	2	Reinforced

The formation of these inhibitory reflexes usually required a great deal of time; although a relative differentiation could sometimes be observed quite early, absolute differentiation was obtained in extreme cases only after more than one hundred repetitions without reinforcement.[2]

How is this obvious inadequacy of the primary phase of the conditioned reflex met? As in the case of the primary redintegrative tendency, a corrective appears in the corresponding inhibitory phase, *i.e.*, in experimental extinction. This was implied in the example just cited.

A special case of this is known in the literature as *conditioned inhibition*. This is of particular interest because it reveals in some detail one of the simplest mechanisms by which sensitivity to the patterning of a stimulus is mediated. Again we choose an example from Pavlov.[3]

A positive conditioned stimulus is firmly established in a dog by means of the usual repetitions with reinforcement. A new stimulus is now occasionally added, and whenever the combination is applied, it is never accompanied by the unconditioned stimulus. In this way the combination is gradually rendered ineffective, so that the conditioned stimulus when applied in combination with the additional stimulus loses its positive effect, although when applied singly and with constant reinforcement it retains its full powers.

[2] I. P. Pavlov, Conditioned reflexes, pp. 146–147.

[3] Idem., *op. cit.*, p. 68.

Irradiation and Differential Inhibition

One of the most clearly marked of the primary tendencies of the conditioning process is that of spontaneous generalization. When a conditioned reflex has been set up in the usual manner, it is found that many other stimuli of a somewhat similar nature will also evoke the response. This is particularly common where a pseudo-conditioned stimulus operates through the same sensory analyzer as the true conditioned stimulus. Under certain circumstances this vicarious spreading of the conditioned tendency may extend even into entirely different sense fields such as from the skin to the senses of the eye and the ear. This primitive tendency to generalization is known among the reflexologists as *irradiation.*

It is evident upon only a little reflection that irradiation is a tendency of enormous importance. Indeed it is hard to conceive how any organism requiring very complex learned adjustments could survive without it. It is a commonplace observation in the animal world that stimuli varying within a rather wide range may require substantially the same reaction. Take, for example, a simple command. Physical analysis of sound shows that the particular stimulus complex constituting a vowel sound such as ä is largely different as spoken by a man and a woman, and even as spoken by the same person at different pitches or different persons of the same sex at the same pitch. Similar variability is found among all sorts of other stimuli which, for most purposes, are considered the same. Indeed it is doubtful whether, in a strict sense, a given stimulus is ever exactly repeated. It follows that if the conditioning process were to be based upon a principle of strictly exact repetitions of the conditioned stimulus, even within the differentiating limits of the analyzer, rarely or never would a sufficient number of such identical repetitions accumulate to raise the conditioning tendency above the functioning threshold. But even if by some miracle of chance a conditioned reflex should get set up under such conditions, of what biological

value would it be? Without the principle of irradiation, it could never function except on the rare chance that the organism should encounter the particular shade of the stimulus upon which the conditioned reflex tendency was originally based. All of the innumerable other shades of variability of the stimulus biologically requiring the reaction could be of no adaptive value to the organism. To be so, each possible shade of the stimulus would need to be separately conditioned. But since the number of such differences would be indefinitely great, the organism might well consume the better part of its life in perfecting the conditioning process of a single response. It is very clear that irradiation is an indispensable principle of learned adjustment.

There is, however, a decided disadvantage in the unlimited tendency to irradiation. If irradiation were extended to its logical limit, it would ultimately bring about a state in which any stimulus whatever would tend to evoke, with little or no distinction, every conditioned response possessed by the organism. This would indeed produce an unadaptive behavior chaos. But, just as we have observed in the two preceding aspects of the primary phase of the conditioning process, an inhibitory tendency enters to save the biological situation. In this third case the corrective tendency is known as *differential inhibition.*

Let us suppose that a conditioned alimentary reflex has been set up to a bell of a certain pitch. Our knowledge of the irradiation tendency makes it quite safe to assume that another bell of a pitch and quality measurably different from the first will also evoke the response. We will assume that the second bell is not a biologically relevant stimulus. In this case it will not, when presented, receive reinforcement. This in turn (assuming an adequately discriminative analyzer mechanism) will gradually develop an inhibition for the pseudo-conditioned stimulus. Meanwhile the true bell will be steadily reinforced which will preserve the biologically valuable conditioned tendency intact. Thus the two tendencies, working jointly, bring about a most excellent adaptation which neither alone could conceivably effect.

The Anticipatory Tendency and Inhibition of Delay

Pavlov describes an experiment[4] in which a dog was given a tactile stimulus continuously for one minute, after which there was a pause of one minute, whereupon some dilute acid was introduced into the dog's mouth. Such an introduction of acid is always followed after a brief interval by a flow of saliva—an unconditioned reflex. Pavlov seems not to have interested himself in the phenomenon here emphasized, so that the detailed timing of the process is not given in his report. It is plain, however, that at the beginning of the experiment the flow of saliva could not have taken place until some seconds *after the termination* of the one-minute pause. After the procedure described above had been repeated a number of times, a significant change takes place. The saliva begins to appear during the one-minute pause *i.e.*, *preceding* the introduction of the acid. The first time there was only half of a drop, presumably appearing just at the close of the period. Ten minutes later, ten drops appear during the pause. Since each drop requires some time for secretion, the first of these drops must have preceded the acid by a considerable part of the one-minute pause. A later repetition yielded fourteen drops during the pause, the first drop of which presumably preceded the introduction of the acid by a still longer interval. This experiment illustrates very nicely a most interesting and significant aspect of the excitatory phase of the conditioned reflex. It is the tendency of the reaction to creep forward in time toward the conditioned stimulus in such a way as to lessen the interval originally separating the two and to make the reaction antedate the presentation of the unconditioned stimulus.

We can now ask what may be the survival value of this anticipatory characteristic of the conditioned reflex. The writer ventures a fairly confident prediction that this primitive mechanism will be found intimately connected with the 'short circuiting' so essential a part of the more complex forms of learning. By *short circuiting* is here meant the tendency of a significant or critical reaction in a learning

[4] I. P. Pavlov, Conditioned reflexes, p. 40.

behavior sequence, to move forward in the series in such a way as to antedate (and thus eliminate) useless and irrelevant behavior segments formerly preceding it. But quite apart from this possibility, the shortening of the time interval between the conditioned stimulus and its response has a most obvious and immediate biological significance. As usual this is most easily seen in the case of defense reactions, particularly those involving withdrawal and flight. If the conditioned defense reaction were to preserve unchanged its temporal distance from the conditioned stimulus, the organism would (assuming the conditioned stimulus to be related in a constant temporal manner to the unconditioned stimulus) encounter the injurious stimulus on every occasion. It would thus in no wise profit by the conditioning of the defense reaction, say, to a distance receptor. This would obviously be very bad biological economy. Clearly, for a defense reaction to be wholly successful, it should result in a complete escape from injury. The only way this can be effected is to have the flight reaction antedate the possibility of the impact of the nocuous stimulus. This the basic anticipatory tendency of the conditioned reflex brings about.

But not all reactions are defensive in this sense. Certain behavior acts such as the various delayed reactions, require for their success in mediating biological adjustment that the period of latency or delay, instead of being reduced to a minimum, shall be separate from the stimulus by a quite definite and fairly prolonged period. This *inhibition of delay*, as it is called, has been studied experimentally by the reflexologists. By special techniques they have been able in dogs to condition periods of delay up to thirty minutes, with considerable precision. These experiments yield convincing evidence that the delay results from an inhibition which represses what would otherwise be an overt tendency for the reaction to follow the conditioned stimulus at once. The following report taken from Pavlov describes one of the more illuminating of these experiments:[5]

[5] *Op. cit.*, p. 41.

The animal can be given food regularly every thirtieth minute, but with the addition, say, of the sound of a metronome a few seconds before the food. The animal is thus stimulated at regular intervals of thirty minutes by a combination of two stimuli, one of which is the time factor and the other the beats of the metronome. Further, if the sound is now applied, not at the thirtieth minute after the preceding feeding, but, say, at the fifth or eighth minute, it entirely fails to produce any alimentary conditioned reflex. If it is applied slightly later, it produces some effect; applied at the twelfth minute the effect is greater; at the twenty-fifth minute greater still. At the thirtieth minute the reaction is of course complete. If the sound is never combined with food except when applied at the full interval, in time it ceases to have any effect even at the twenty-ninth minute and will only produce a reaction at the thirtieth minute—but then a full reaction.

Once more, then, we observe the primary excitatory phase and the secondary inhibitory phase of the conditioned reflex combining in a kind of trial-and-error process to bring about a type of biological adaptation which neither tendency could possibly produce alone.[6] The tentative or trial process is mediated by the excitatory phase; the selective or corrective process is effected by the inhibitory phase—at bottom, experimental extinction brought about automatically by failure of reinforcement.

The Dilemma of the Conditioned Defense Reaction

In connection with that aspect of the conditioned reflex last considered, a curious and rather sharp distinction appears between positive reactions such as those involved in the taking of food, and defense reactions such as involve withdrawal or flight. In the case of an alimentary reaction, a successful response would ordinarily be followed each time by the consumption of food. This means, of course, that the conditioned tendency is continuously reinforced, which will keep it up to full strength. In this respect the case of the defense reaction is quite otherwise. As pointed out above,

[6] While undoubtedly related, this process is not to be confused with the trial-and-error of ordinary learning such as of the maze. The author hopes in a later paper to elaborate this distinction.

for a defense reaction to be wholly successful, it should take place so early that the organism will completely escape injury, *i.e.*, the impact of the nocuous (unconditioned) stimulus. But in case the unconditioned stimulus fails to impinge upon the organism, there will be no reinforcement of the conditioned tendency which means (one would expect) that experimental extinction will set in at once. This will rapidly render the conditioned reflex impotent which, in turn, will expose the organism to the original injury. This will initiate a second cycle substantially like the first which will be followed by another and another indefinitely, a series of successful escapes always alternating with a series of injuries. From a biological point of view, the picture emerging from the above theoretical considerations is decidedly not an attractive one.

The sharpness of the conflict here, invites speculation as to how the problem is met by nature. One possibility which suggests itself is that the greater potency of the defense reaction tendencies may make them less subject to the weakening tendencies of experimental extinction. Another possibility is that the tendency to experimental extinction may be more or less in abeyance where defense reactions are concerned. But as soon as the principle of experimental extinction becomes inoperative, the organism is exposed to the dangers resulting from accidentally conditioned irrelevant stimuli (p. 501). There is thus presented a kind of biological dilemma apparently not at all the product of misplaced ingenuity on the part of the theorist. If experimental extinction operates fully the organism seems doomed to suffer the injury of the nocuous stimulus periodically in order to renew the strength of its conditioned defense reactions. If, on the other hand, experimental extinction does not operate, the organism seems doomed to dissipate much of its energy reacting defensively to irrelevant stimuli.

It is suggested on the basis of mere casual observation that what might be called a kind of organic compromise may be operating in this curious situation. It may be that experimental extinction becomes progressively in abeyance as

the gravity of the injury increases. Thus slight injuries would suffer considerably from experimental extinction and would consequently require more frequent nocuous reinforcement. Reactions to grave injuries would be affected relatively little by experimental extinction but for this reason would be very prone to become attached to irrelevant stimuli. This last, indeed, may account for the prevalence of phobias which appear, at least superficially, to be more or less accidental conditionings of irrelevant stimuli to strong emotional reactions. On the other hand very mild punishment is very likely to require frequent repetition. The problem presents a fascinating field for experimental investigation.

[MS. received April 20, 1929]

BEATS AND RELATED PHENOMENA RESULTING FROM THE SIMULTANEOUS SOUNDING OF TWO TONES—II

BY ERNEST GLEN WEVER

Princeton University

The Beats of Mistuned Consonances

If two tones, one the exact octave of the other, are sounded together, we hear a fused, continuous tone; but if one of the tones is altered slightly in frequency we hear beats. Likewise, beats arise from other consonant, just-tempered intervals: fifths, thirds, double-octaves, etc.; when they are mistuned. These are the beats of mistuned consonances, as distinct from the beats previously accounted for in this paper.*

Historically, the problem of the nature and origin of these beats is as old as that of the more familiar sort. Both problems arose side by side in the field of musical harmony. Since in concords played on musical instruments the 'beats of mistuned consonances' are only a little less prominent than the 'beats of imperfect unisons,' it is natural that the former, no less than the latter, should have played a prominent rôle in musical theory.

But it should be said at once that so far as the auditory experience is concerned these beats do not in any essential respect differ from the beats earlier discussed; they are likewise perceived as variations in the intensity of tones. The problem which these beats bring is the problem of their origin.

There are three theories as to the origin of the beats of mistuned consonances. These I shall designate as (1) the overtone theory, (2) the transformation theory, and (3) the resultant displacements theory.

* Part I of this article appeared in the preceding number of this Journal: Psychol. Rev., 1929, 36, 402–418.

The Overtone Theory—According to this theory the notes sounded contain overtones, and the beating involves these, and not merely the fundamentals. Beats arise in the perception of two of these components whenever their frequencies are sufficiently close together.

Suppose, for example, that we sound c (256) together with c' (512); there would be no beats. But if the former note were mistuned, say, to 255, we should hear beats; and according to the theory they arise from the first overtone of the lower note ($2 \times 255 = 510$) and the fundamental of the higher (512). The calculation ($512 - 510$) gives 2 beats per sec., which agrees with experience.

Beats from intervals other than the octave are accounted for in like manner. One more example will suffice. Suppose we mistune the upper note of a fifth, the interval between the tones 200 and 300, so that we obtain the tones 200 and 302. The second overtone of the lower note will be 600, and the first overtone of the higher will be 604; these will beat at the rate of 4/sec., as experience verifies.

This method always gives the proper number of beats; but then the same is true of the other two theories soon to be discussed.[41] The calculation will be simplified by use of the following formula:

$$mN - nM = b, \text{ the number of beats,}$$

where m/n is the simplest perfect ratio of the interval, and M, N are the corresponding vibration-numbers of the tones which beat. Illustrating with the mistuned fifth, as in the example above, $m = 2$, $n = 3$; $M = 200$, $N = 302$; hence,

$$2 \times 302 - 3 \times 200 = 604 - 600 = 4 \text{ beats/sec.}$$

[41] Hermann denies this for the resultant displacements theory: see *Arch. f. d. ges. Physiol.*, 1894, **56**, 483 *f.*; *Ann. d. Phys.* (new series), 1896, **58**, 392 *f.* But his argument is not convincing to me; he simply infers from the forms of the curves that if phase made a difference we should hear twice as many beats as we do. Koenig's results with his wave siren seem to me to cut the ground from under this position; see Koenig's Quelques expériences d'acoustique, 1882, No. 16, esp. p. 229.

This formula corresponds to that developed by Smith, but is rather simpler.[42]

It is evident that the overtone theory places the beat of mistuned consonances on the same plane with the unison beat, for the beating is always regarded as due to the interaction of two adjacent frequencies simultaneously acting on the ear. And hence the primary validity of the theory rests upon the question of the physical presence of overtones in the sounds used. Hermann argued that such was the case with the octave beat.[43] And, indeed, it must be admitted that the presence of overtones is the rule in musical sounds, and the theory holds, in part at least, for the explanation of the beats of musical notes.[44] But the general validity of the theory depends upon whether beats arise when the tones are pure.

That beats do arise under these conditions was maintained by Koenig.[45] He used tuning-forks whose waves he considered to be free of overtones, and obtained powerful beats from a great number of mistuned intervals. Indeed, he was able to perceive beats distinctly at the triple octave ($c : c^3$, ratio 1 : 8), which on the overtone theory would require the appreciable presence of the seventh overtone of the lower note. And this he was unwilling to admit.

The recent work of Wegel and Lane is in confirmation of Koenig. They obtained beats of higher orders from tones which they showed to be practically pure. Their tones were produced by audio-oscillators and were passed through electric filters, and our knowledge of the adequacy of such filters gives strong credence to their claim that harmonics were sufficiently eliminated. Moreover, they used a condenser-transmitter to pick up the sounds from the air, and were able

[42] See De Morgan, On the beats of imperfect consonances (1857), *Trans. Cambridge Philos. Soc.*, 1864, **10**, 129–141. Smith's formula used interval units instead of vibration-frequencies; see Harmonics, 2d ed., pp. 82 *f.*

[43] L. Hermann, Zur Frage betreffend den Einfluss der Phasen auf die Klangfarbe, *Ann. d. Phys.* (new series), 1896, **58**, 393 *f.* Helmholtz seemed of the opinion that upper partials account for Koenig's results; see Sensations of Tone, p. 159.

[44] Helmholtz, *op. cit.*, p. 180.

[45] Koenig, Ueber den Zusammenklang zweier Töne, *Ann. d. Phys.* (2d series), 1876, **157**, 177 *ff.*

thus to limit the further possibility that partials were produced in the external medium.[46]

In the presence of this evidence I think we can safely conclude that the overtone theory must be ruled out as a complete explanation of the beats of mistuned consonances. Though it no doubt holds for most musical tones, it breaks down with the use of pure tones, since with these the beats are still found to occur.

The Transformation Theory—This theory differs from the preceding in the assumption that the beats arise, not from overtones actually present in the sounds as they reach the ear, but from tones which have their origin in the auditory mechanism itself, through a transformation of part of the energy of the primaries into periodicities of higher orders. On this theory a pure tone, when it is strong, will give rise to a sort of Fourier's series in the ear.[47]

Transformation occurs in a conducting system when the response is not a linear function of the force applied. The supposition of transformation in the conducting mechanism of the ear is entirely reasonable in view of the peculiar conformation and weighting of the ear-drum, and the apparent loose coupling of the ossicles; for these make it likely that the displacement of the stapes is not at every moment strictly proportional to the amplitude of the wave by which it is actuated.[48]

The tones resulting from the transformation (which may be either overtones or combination-tones) are then considered to play the same rôle as is assumed for true overtones in the first theory.

This theory is to be found in Helmholtz alongside the overtone theory. Bosanquet stated the theory clearly in 1881, and defended his position against the attack of Koenig.[49]

[46] Wegel and Lane, *Phys. Rev.*, 1924, **23**, 272 *f.* For the production of partials in the air, see Helmholtz, Ueber Combinationstöne, *Ann. d. Phys.* (2d series), 1856, **99**, 497 *ff.*

[47] *Cf.* Wegel and Lane, *op. cit.*, 273 *f.*

[48] Helmholtz, Sensations of Tone, p. 158.

[49] Bosanquet, *Phil. Mag.* (5th series), 1881, **11**, pp. 426, 499 *ff.* See also his article, On the history of the theory of the beats of mistuned consonances, *ibid.*, **12**, 275 *ff.*

Koenig took the position that we cannot assume the interaction of tones whose presence themselves cannot be demonstrated. Because he was unable in the complexes of sound to hear the overtones and combination-tones that the theory presumes, he concluded that they are not there. He said, "One is obliged always to assume the existence of tones which are not only themselves unheard, but are often also both the result of certain tones, and the cause of others, which are all equally inaudible." [50]

This is a reasonable-sounding argument; but it cannot be admitted. For while these tones may not be present in perception, they may nevertheless be present in a physiological sense in the peripheral mechanism, and when properly reinforced by adjacent periodicities may manifest themselves in perception. For, as shown above, beats may indeed arise from (physical) tones which are too low in intensity to be heard even when presented alone. And it has been shown further that such a tone may exist in a complex in considerable intensity, far above the threshold of the tone when sounded alone, without being detectable in any other way than through its beats; because it is to a high degree masked by the other members of the complex.[51] Hence, contrary to Koenig, we must admit the possibility of the presence, in the receptive mechanism, of overtones and combination-tones such as the transformation theory requires, even when these tones are not perceptible as such.

Two important observations may be mentioned here. The first is, as noted by Koenig, and later by Bosanquet, that typically the beats of mistuned consonances are perceived as variations of the lower note.[52] This may be explained in terms of the phenomenon of masking, for it is known that low tones tend to obliterate high tones, but that the reverse is true only to a slight extent. What happens, then, in the case of the mistuned consonance, is that the lower note tends to mask the higher, so that the higher is less prominent in the ex-

[50] Koenig, *Ann. d. Phys.* (2d series), 1876, **157**, 200*f.*

[51] See Part I, PSYCHOL. REV., 1929, **36**, 408*f.*

[52] Koenig, *op. cit.*, p. 188; and Bosanquet, *Phil. Mag.*, 1881, **11**, 427 *ff.*

perience, yet is present to a sufficient extent to produce beats. A second observation is related to this one, the observation that the best beats of mistuned consonances are obtained, not with equal intensities of the two tones, but when the lower note is the stronger.[53] This fact lends itself readily to explanation in terms of either of the above two theories, for both theories maintain that the effective periodicity is of a higher order for the lower tone than for the higher. Thus, *e.g.*, in the case of the octave, the first overtone of the lower note is supposed to beat with the fundamental of the higher, and hence the lower note must be made stronger so that the first overtone may become effective.

The Resultant Displacements Theory—This third theory has nothing to do with overtones or combination-tones, however produced, but attempts to derive the beats from the primaries alone. It supposes that the two periodic forces are exhibited as resultant displacements in a common receptive mechanism. Or in other words, what we hear corresponds to such a wave as we would obtain by adding the amplitudes at all points, and the beats are represented by regions of maximum amplitude in the resultant.

This theory is the oldest of the three, going back for its beginnings to Smith's classical treatise in 1749; for, indeed, it was Smith's peculiar contribution to the whole subject to point out that in the geometrical forms representing imperfect harmonies there are certain patterns of resultant displacements that recur.[54] And the periods of these patterns correspond to those of the beats under consideration.[55]

Thomas Young is usually regarded as having held to this theory, though he paid no debt to Smith, and gave the theory but slender formulation.[56] Koenig stated the theory definitely.[57] He said:

[53] W. Thomson, On beats of imperfect harmonies, *Proc. Royal Soc. Edinburgh*, 1877–8, **9**, p. 610; Hermann, *Ann. d. Phys.* (new series), 1896, **58**, p. 393.

[54] Robert Smith, Harmonics, 1759, 2d ed., 56 *ff.*

[55] See, however, note 41.

[56] Young, Course of lectures, 1807, I, 387 *f.*, 389 *ff.*

[57] Koenig, *Ann. d. Phys.* (2d series), 1876, **157**, 186 *f.*

It is therefore far more natural to derive the beats of the harmonic interval, like those of the unison, directly from the composition of the sound waves, and to consider that they arise from the periodically interchanging coincidences of the similar maxima of the primary tones, and of the maxima which have contrary signs.

As is evident in this quotation, the principal argument for the theory is based upon its simplicity and directness. These are desirable qualities in a theory, and were no further feature presented for consideration, we should be warranted in accepting this theory in preference to the transformation theory, which admittedly involves the further assumption of transformation. But there is some doubt about the simplicity of the resultant displacements theory. We are not told much about the actual workings of the theory, beyond the hint that the ear perceives the 'resultant forms.' But these resultant forms are themselves by no means simple. If we consider some of the curves drawn to represent mistuned consonances [58] we find no gross variation in amplitude, such as occurs for the unison beats. True, the eye can analyze out of the curves certain wave-patterns which correspond to the beats; but such visual analysis suggests the necessity of some corresponding analysis by the ear when the complex is presented in audition. Yet we have no indication in Koenig's theory, or elsewhere, as to how this takes place.

And in the resultant displacements theory, as Bosanquet has pointed out,[59] a further matter is involved. This theory, in maintaining that tones widely separated in frequency produce beats by acting simultaneously on a common receptor, is running in the face of Ohm's law, for if we have peripheral analysis it is difficult at the same time to conceive of peripheral interaction. The transformation theory, on the other hand, no more violates Ohm's law than does a theory of ordinary beats, for it holds that the consonance beats, like the simple beats, are ultimately produced by tones of adjacent frequency. Consequently, if we accept Ohm's law, we are led

[58] Many have prepared and published such curves. Some of the best are exhibited in Koenig's Quelques expériences d'acoustique, 1882, p. 97.

[59] Bosanquet, *Phil. Mag.*, 1881, 11, 422 *f.*

away from an acceptance of the resultant displacements theory, and toward the theory of transformation.

Beats and Auditory Theory

A consideration of beat-phenomena seems to demand our acceptance of Ohm's law as valid except for tones which fall within a limited range of frequency. In our search for an adequate theory of hearing, therefore, we must be guided by the necessity of explaining not merely analysis as such, but analysis of a particular, limited kind: so limited that on the same theory we can account for the production of beats.

Of auditory theories, as everyone knows, we have a great many. Most formulations of these theories include an explanation of beats; and probably all of the more important theories would allow such explanation. But the Helmholtzian resonance-place type of theory has been carried farther toward a complete explanation of beat-phenomena than has any other; and hence when confronted with the necessity of making a choice, I have taken this type of theory as a basis for the discussion which follows.

In adopting a resonance theory here, however, particular attention must be directed to the sort of resonance that is hypothesized. The concept of perfect resonance will not serve, for a perfect resonator will respond only to the frequency to which it is tuned, and if the ear were fitted with a system of such resonators we should have complete analysis, and no beats. But a resonator sufficiently damped will respond in greater or less degree to a whole band of frequencies, and this is the sort of mechanism that the situation requires. For with such a system the ear could act as an analyzer when tones were remote in frequency, and as a synthesizer, giving beats, when the frequencies were adjacent.[60]

With the consequences of this theory for hearing in general I shall not at this time be concerned; I shall turn rather to a consideration of the theory as an explanation of the phenomena of beats as set forth in this paper.

[60] *Cf.* H. E. Roaf, The analysis of sound waves by the cochlea, *Phil. Mag.* (6th series), 1922, 43, 349–354; and H. Fletcher, *Bell System Tech. J.*, 1923, 2, No. 4, 165 *ff.*

Analysis and the Beating Complex—On this theory, as has been supposed, analysis will not occur when two tones are so close in pitch as to involve considerably overlapping bands of the resonant membrane; here the response of the membrane may be represented as the resultant of the primary responses. When the primaries are in phase their effects combine on those parts of the membrane common to the two to produce a maximum displacement; and when they are in contrary phase their effects interfere to produce a minimum displacement. And these cycles of maximum and minimum displacement are perceived as periodic variations of intensity, which constitute the beats. The periodic changes involve both amplitude of displacement and extent or spread of stimulation on the membrane; the one may mean greater frequency of impulses in the stimulated nerve-fibers, the other a greater number of fibers excited; and either or both these influences we may regard as bringing about the perception of greater intensity. In this case but one tone will be heard, of a pitch intermediate between the primaries, corresponding to that position on the membrane where the disturbance is greatest; this assumes Gray's principle of maximum stimulation.[61]

At a certain frequency analysis sets in, but is incomplete, and the primaries are heard in addition to the intertone. Here we may suppose that at coincidence of phase the overlapping of the primary bands is such as not to involve a definite intermediate maximum as in the preceding case, but a broad band over which the stimulation is about the same, while at opposition of phase there is a minimum of disturbance in the intermediate region, and two maxima corresponding roughly to the primaries on either side. Hence in perception we might have the primary disturbances pretty constantly represented, with the intertone coming in when these summate at their coincidence of phase.

At a further separation, analysis is practically complete, and only the primaries are heard, slightly beating. Here the overlapping is but slight, and the summational effect is never so great as to create an intermediate maximum of stimulation.

[61] G. Wilkinson and A. A. Gray, The Mechanism of the Cochlea, 1924, 152 *ff.*

Yet the phase-variations give regular cycles of displacement within the region of overlapping, sufficient to produce noticeable changes of intensity, and hence beats.

Further separation gives two tones, without beats; here analysis is perfect. The overlapping is absent, or too slight to give any detectable variation of intensity.[62]

Beat-Thresholds—The absence of any lower threshold of beat-frequency, as I have suggested is the case, is in accord with the conception of the peripheral mechanism as here outlined. However slow the change of phase, we should always get excitation at coincidence and minimum or zero excitation at opposition.

It is reasonable to suppose that the upper threshold of frequency for beat-perception is imposed by the inertia of the resonant membrane, for the oscillation between maximum and minimum of the beat necessitates a shearing type of movement of the membrane in addition to its vibratory movement; and to explain the variation of the threshold with pitch-region we have only to consider that the inertia is greater at the lower regions of pitch. These considerations are quite in line with the theory in question; for just such a gradation of inertia is already assumed by that theory in its explanation of differential response as regards the frequency-range.

The theory in respect to intensity is simple enough. The loudness of the beat at maximum represents the combined energies of the two waves in their effect upon the resonant membrane; while the loudness at minimum represents the difference in these energies when the waves are opposed. The loudness thus varies from a maximum that is above the loudness of either component, to a minimum that under proper conditions is silence. This conception makes it clear why we are able to perceive beats when one of the components is at a subliminal intensity: while it alone is not audible, in the combination it is nevertheless effective in producing noticeable variations of intensity. Likewise, one should hear beats from two tones separately inaudible (though not far below the

[62] *Cf.* Stumpf, Tonpsychologie, II, 484 *ff.*

threshold), if during the periods of coincidence of phase in the beat-cycle their added amplitudes were sufficient to cause supraliminal displacements of the membrane.[63]

The Beats of Mistuned Consonances—If the transformation theory is accepted, no extension of the general theory, beyond the matter of transformation itself, is required to explain the beats of mistuned consonances, for these beats are regarded as arising actually from tones of adjacent frequency.

Summary

In this paper is reviewed our knowledge of the phenomena of beats: the conditions under which they arise, their subjective nature, and their relation to auditory theory.

1. Beats arise from the coincident action of two tones of adjacent frequency. As the frequency-difference of the tones is raised from zero upward, three stages of the phenomenon are distinguished: (1) a stage of noticeable surges of intensity, (2) a stage of intermittence or pulsation, and (3) a stage of roughness without intermittence.

2. The thresholds of beat-perception are shown to vary with (1) the pitch-region from which the tones are taken, (2) the intensities of the tones, and (3) the criterion of judgment. Existing data are reviewed regarding the upper limit of frequency for beat-perception; but it is probable that no true lower frequency-threshold can be established. The limits of intensity are defined, and the relevance of the phenomenon of masking indicated; the complete possibilities of beat-perception for a pair of tones are pointed out in a single instance.

3. Under the heading of the beating complex are discussed in detail the nature of the auditory experience in the presence of beats, and the variation of that experience as a function of the frequency-difference of the beating tones.

4. The beats of mistuned consonances are considered, and the evidence weighed with respect to the three theories of

[63] Wegel and Lane, *Phys. Rev.*, 1924, 23, p. 270.

their origin: (1) the overtone theory, (2) the transformation theory, and (3) the resultant displacements theory.

5. Finally, with a resonance theory of audition taken as a working basis, a theoretical account is attempted for the phenomena exhibited, particular regard being given to Ohm's law and the nature of the beating complex.

[MS. received March 25, 1929]

CRITIQUE OF THE SEASHORE CONSONANCE TEST: A REPLY TO DR. LARSON

BY CHRISTIAN PAUL HEINLEIN

Florida State College for Women

In the 'Psychology of Musical Talent' (**17**), C. E. Seashore has unequivocally declared that musical talent, which he holds to be a hierarchy of talents, is an inborn gift, and that the musical mind is distinguished by specific traits or factors which are essential for the apprehension and expression of certain attributes of sound, namely, pitch, intensity, duration, and extensity. Among the significant factors of the musical mind is listed 'musical sensitivity,' which comprises a group of eight native capacities or 'senses' manifested as either simple forms of impression or complex forms of appreciation (**17a**). Included in the complex forms of musical appreciation is the 'sense' of consonance, which is defined (**17b**) as 'the basic power that underlies the ability to judge æsthetic effects in combinations of tones.' The Seashore Consonance Test (**17c**) which has been devised for the purpose of obtaining an objective index of an individual's discriminatory precision in an æsthetic judgment of this type is so well known to psychologists and musical educators alike that any detailed description of its structure is a needless introduction.

In 1925, Joseph Peterson (**16**), in discussing the functional aspects of consonance, pointed out that tonic effects, anticipated resolutions, and harmonic associations which resulted from the paired comparison method employed in the Seashore Consonance Test greatly affected the judgments of the listeners. Following this discovery, Peterson indicated doubt as to whether there really existed 'a purely independent comparison of two sets of intervals on their own degree of consonance.' In the same year, the writer (**5**) conducted an experimental investigation on the Seashore Consonance Test, obtaining results which emphatically confirmed the

results presented by Peterson. It was shown that the test gave not only a low coefficient of reliability—a feature further corroborated by the recent findings of Brown (**2**) and Brennen (**1**)—but that paired interval comparison is an inadequate method for testing consonance, since such comparison through contrast of clang richness and through recognition of harmonic relationship is conducive to a strong affective bias in judgment.

Dr. Delia Louise Larson (**11**) has recently published a monograph in which she gives an experimental critique of the Seashore Consonance Test. She tested six groups of subjects, both musically trained and untrained, and from an analytic study of the data obtained concludes that the paired-interval comparison method is adequate for testing consonance and that the principle of harmonic progression did not function to distort the scores. Thus it is immediately apparent that the results obtained by Dr. Larson and those obtained by the writer stand at opposite poles. Dr. Larson has introduced the criticism that in my instructions to the subjects I failed to counteract the tendency to judge consonance on the basis of pleasantness and unpleasantness, and that 'preference' was not eliminated as a criterion. Since this criticism centers about the nature of the test instructions, let us pause and analyze in detail the content of these instructions, and the theoretical considerations which have entered into their adoption.

In order to understand clearly the development of the present consonance test and the use of the several criteria which are employed in the instructions, it is necessary to turn back to Malmberg's article (**13**) on the perception of consonance and dissonance which gives us an interesting historical survey of the various theories and definitions of consonance. After a brief resumé of many antagonistic opinions as to what constitutes consonance, Malmberg recognizes the following criteria as indicative of consonance: *blending*, *smoothness*, *fusion*, and *purity*—criteria which have been borrowed from Helmholtz (**8**), Stumpf (**19**), Wundt (**20**), and Krueger (**10**). In a revision of Malmberg's test,

Gaw (4) omits the factor of 'purity'—thinness of tone, absence of richness—as a criterion of consonance. It is this later revision which Seashore has followed very closely.

First of all, let us note the criteria of consonance which Seashore has adopted. There are three—*smoothness*, *blending*, *fusion*. These are the criteria which are included in the 'Manual of Instructions' (18). But these criteria do not agree with the criteria which Seashore adopts in his 'Psychology of Musical Talent.' Gaw borrowed the criterion of *fusion* from Stumpf, whereas Malmberg, borrowing from Krueger, substituted the criterion of *purity* instead. In the 'Psychology of Musical Talent' (17d), Seashore in following Malmberg adopts *blending*, *smoothness*, and *purity* as criteria of consonance, omitting *fusion* because "it does not agree with the ranking in the three criteria here adopted." *Fusion* is furthermore left out of account because "it means more than blending," because "it represents the factor of unanalyzability which, in some cases, contradicts the other criteria." Malmberg states that he omitted *fusion* as a criterion because "it resulted in a peculiar classification of its own." Notwithstanding these objectionable features, *fusion* is later introduced as one of the three significant criteria (18). This would lead one to believe that the originators of the test material were not certain just what did constitute consonance. At any rate, whether the criteria are fundamental or superficial, whether their choice was arbitrary or natural, the fact remains that the adopted concept of consonance is determined and delimited by these chosen criteria.

At most, we can say that the criteria as presented are both relative and indefinite. *Blending*, *smoothness*, *purity*, and *fusion* are relative terms potent to arouse various types of affective associations in the minds of the listeners and thereby coloring judgment of the tonal clangs. It must be remembered that these terms are borrowed from non-auditory experience, and hence tend to introduce any number of associations relative to non-auditory experience of various kinds and degrees. Moreover, for auditory experience, each

criterion itself demands further definition. After reading the instructions to highly intelligent senior college students, I have been frequently asked: "What do you mean by *blending* —by *smoothness*—by *fusion?*" Children unfamiliar with this terminology usually manifest a blank stare of apparent confusion or else they tend to indicate indifference to the instructions which they are unable to comprehend. To subjects unsophisticated in the abstruse speculation of Stumpf, the term *fusion* means little or nothing.

Assuming offhand that the listener gains some notion of what the test criteria might signify, consider for a moment the method of comparison which he is instructed to employ. In each set of paired intervals, each interval is sustained for approximately two seconds; between the intervals of each pair there is a rest of one second. According to Seashore's musical notation, a rest of two seconds only is allowed between each pair. During this very brief sustension of the interval, and during the exceptionally brief pauses, the reactor is expected to carry in mind three criteria relative in meaning (in fact, six criteria when the opposites are taken into account), and to subject each interval, through careful comparison, to an agreement or disagreement with the three proposed criteria (**17e**). Such a *tri-dimensional judgment* is to be continued for fifty pairs. It seems impossible for even the most intelligent reactor to carry out such analytic comparison *in toto;* in view of the very short time allotted, the impracticability of the procedure is at once obvious.

Seashore, in allowing his listeners preliminary practice, follows closely in the footsteps of Malmberg, who illustrated his criteria by presenting to his subjects simple intervals and explaining their consonance value. The danger of such preliminary illustration and explanation can be seen at once. A somewhat analogous situation would be one in which the subject were actually told that a particular color presented to him was red, and a few moments later, upon presentation of the same color, the subject were asked to name the color. First the observer is tempted to sin and thereafter forbidden to do so. It is the old story of 'trying hard to forget.'

How is one able to prevent the listener from memorizing the values thus attached to these preliminary illustrations, and how is one able to ascertain that such memorization has or has not taken place? It is reasonable to suppose that any intelligent observer would through the function of common-sense naturally incline to attach these designated values to the same intervals upon their recurrence in the test proper.

Seashore, apparently, does not discriminate between 'practice' and 'training.' When a portion of the test proper is employed in the preliminary explanations—such as the use of the A-side of the phonographic disc which a few moments later becomes an integral part of the experiment—the preliminary period should be properly interpreted as a period of *training* rather than one of 'practice.' According to the methods and principles of scientific psychology (**3**), under such conditions as those cited above, comparison of results obtained by different investigators becomes a worthless procedure. Nor is it proper to compare results obtained under such varied conditions with some pre-established norm. The prime requisite in mental measurements of this type is constancy of conditions.

We must not forget, however, that according to Seashore the sense of consonance is an inborn trait, a native capacity of the organism, one of the most fundamental and essential capacities of the musical mind. If this is true, then any type of preliminary training and education will never reveal the innate value of such a capacity. Seashore, however, has never been able to prove that the sense of consonance *is* an inborn capacity, nor has he ever demonstrated that the sense of consonance such as he understands it to be is a fundamental or an essential capacity of the musical mind. Such assertions are based on purely *a priori* assumptions. On the basis of recent genetic discoveries (**9**), the scientific psychologist can no longer maintain that it is necessary to have a certain characteristic because one has inherited a certain set of genes. Nor is it true that the appearance of certain traits signifies dependence upon the direct transmission of specific genes possessing specific potentialities. It is well

known that under different conditions a given set of chromosomal factors will produce very different results. If a listener scores a high grade in the consonance test after preliminary 'practice,' his score is no certain indication that he has inherited a set of genes whose function favors this result.[1]

In a summary analysis (**17f**) of the sense of consonance as a native capacity, Seashore emphatically declares that consonance is ultimately feeling. In the light of this admission, it is indeed exceedingly difficult to understand just why Seashore rules out feeling-tone from the perception of consonance. One moment we are told that feeling-tone is an integral part of every consonance experience, and in the next moment we are told that consonance perception is a cognitive act of discrimination in which feeling-tone and association are eliminated (**17g**).

Dr. Larson (**11b**) contends that the judgments of my subjects were affectively colored according to the nature of the introspections which they gave. I may suggest that it was my purpose to present these several introspections in order to demonstrate to the reader the natural prevalence of very strong feeling-tone in judgments of this type. I have never denied that the judgments of my subjects were affectively colored: on the contrary, I have insisted that affective coloring was an inescapable condition in consonance experience (**5a**). Dr. Larson further explains that affective rather

[1] Larson (**11a**) admits that "a high score in the Seashore Consonance Test is not necessarily indicative of musical capacity."

The science of genetics has failed to demonstrate the existence of any hereditary vehicle which might bear a unitary intellectual talent. In the main, biometric analyses, such as those conducted by Stanton, Hurst, Ziehen, and others, which depend upon correlation ratios to demonstrate familial similarities in mental abilities neglect to a great extent the functional importance of environmental conditions in moulding the life-career of the individual. Uninformed writers on the subject of heredity have endeavored to impress the popular mind with the notion that the Mendelian laws of inheritance are applicable to the complex intellectual talents of man as well as to the physical characteristics of the lower forms of insect and animal life. Intellectual talents considered as dispositional aspects of neurological functioning are not inherited in the same way that physical characteristics are inherited. The search for 'musical' chromosomes is analogous to the search for 'soul' chromosomes.

An excellent impartial discussion on the effects of nature and nurture on musicality will be found in the Twenty-seventh Yearbook of the National Society for the Study of Education, 1928, Part 2, by Paul R. Farnsworth.

than cognitive judgments resulted because no preliminary practice was afforded. This, indeed, is just what I had hoped to prove. I deliberately avoided preliminary practice in order to demonstrate that in a natural, untrained judgment of consonance according to the criteria presented, feeling-tone and musical associations were integral aspects of nearly every consonance experience (**5b**). It is a relatively simple matter to demonstrate an affective bias in binary clang experience, but it is an extremely difficult proposition to demonstrate that a condition biologically fundamental is eliminated through any naïve instruction to effect emotional repression.

Practically every investigation that has been conducted on consonance discrimination has suffered the persistent invasion of feeling-tone and musical association. Dr. Larson admits of the difficulties encountered in efforts to counteract the tendency to make judgments on the basis of preference.[2] Her criticism is indeed well founded in regard to the affective connotation of the terminology introduced into the test instructions. We are told by Seashore that 'harmony' is a word which indicates a strong affective experience (**17f**), and yet this very term is included in the instructions. It is insisted that every effort should be made to eliminate affective experience in judgment, yet the terms 'good,' 'better,' 'bad,' and 'worse,' which imply hedonistic preference, are included in the instructions. Furthermore, several marked ambiguities arise from the use of this terminology. A *good* combination is explained on the basis of particular criteria. A *bad* combination is simply referred to as an 'opposite' one. Thus we have certain 'standards' established for comparison.

[2] Larson (**11c**) writes that "consonance being allied to perceptual processes, involves imaginal as well as sensory factors: not only are the sensory elements to be noted, but the imaginal factors that hark back to the individual's previous experience and to his contacts with the traditional standards of the race must be considered. This brings us directly to the type of judgment involved in the matter of consonant and dissonant intervals. It appears at once that since there are both sensory elements and imaginal background involved, the judgment calling for a fusion or 'oneness' of the sensory blend may be influenced by traditional associations of an æsthetic nature. More specifically, too, the sounds produced will be colored also by affective processes."

Following this establishment of a 'type,' we are asked to judge in terms of *degree*. Confusion in judgment is bound to follow when two *good* combinations or two *bad* combinations represent a single pair. For example, let us suppose that a single pair comprises the intervals of a major Third and a perfect Fourth in the order named. Both intervals, according to definition, are *good* combinations. Let us assume in agreement with Seashore that the first interval is the *better* combination. Then, according to the prescribed use of the terms *better* and *worse*, we must consider the second interval *not* as a *bad* combination, but as a *worse-good* combination.

But if the perception of consonance as a cognitive act of discrimination becomes affectively colored, what might we expect in so far as test results are concerned? The answer is—*negative results*. Dr. Larson, in support of certain findings in my previous investigation, tells of a subject with eight years of musical training scoring 88 and 90 in her tests who later declared: "I took this same test two years ago and made a score of 46, but my judgments at that time were made solely on the basis of pleasantness and unpleasantness." This is quite in keeping with my own findings (**5c**), in which musically trained subjects, without being given the opportunity to learn in advance the proposed consonance values of certain intervals through test demonstration, made relatively low scores, and untrained subjects, whose affective appreciation of binary clangs is generally low, made relatively high scores.

Malmberg (**13**), in referring to the definition of consonance which he had adopted, seemed to feel quite confident that agreeableness, which had played such an important rôle in theory, was here conspicuous by its absence. His observers were instructed to judge each binary clang as an æsthetic object by itself, without respect to the effect of progression, meaning, association, or mood. Yet in the data which he obtained, he discovered that the Third was 'unduly preferred' by the reactors. He admitted that the factor of agreeableness had not been satisfactorily eliminated. He declared that "the *musical value* of the Third as compared with the Octave must have influenced the students to some extent."

Gaw (4) likewise found that the major Third was preferred to the Octave as the 'best' consonance. Instead of fully acknowledging the natural and fundamental rôle which feeling-tone plays in practically every consonance judgment in which clangs are contrasted, Gaw subordinates this primary psychological process by contending that it is 'the mediocre and poor observers' who are thus influenced by this richness in one of the members of the pair of intervals. Dr. Larson offers a similar explanation, adding, moreover, that undue preference for the rich combinations is due to 'lapses in attention.'

Malmberg himself (13a), in statements together with ample data presented, furnished marked evidence in favor of the fact that musically inclined subjects find it extremely difficult to suppress feeling and association in binary contrast. Much to his indicated surprise, he discovered from a large amount of data which he collected that the coefficient of correlation between musical status (which included musical training, musical ability, musical environment, and the expression of musical feelings) and the record for the Seashore Consonance Test is practically zero—only 0.02. Malmberg remarks that such a low correlation 'is very remarkable in view of the fact that one should ordinarily expect those who had musical training to do better than those who had none.' His findings in this respect have been confirmed by Brennen (1a), who also obtained coefficients of 0.02, P.E. .17 between musical training and the record for the consonance test, and 0.28, P.E. .14 between musical capacity and the record for the consonance test. My own findings confirm both of these investigations. A chief point of disagreement between Dr. Larson and myself apparently centers about the question as to what constitutes a significant correlation and coefficient of reliability. Assuming a low probable error, I do not consider a coefficient of 0.60 as being very significant. I should never consider a test very reliable which did not give a coefficient of at least 0.80. Brown [3] obtained a reliability

[3] Brown (2a) obtained a correlation coefficient of 0.17, P.E. .06 between consonance and teachers' judgment of musical ability.

coefficient of 0.43 in a repetition of the test, while Dr. Larson (**11d**) obtained reliability coefficients of 0.63 and 0.65.

Despite the fact that the test originators have recognized the persistent prevalence of feeling and association in judgments of consonance, every attempt has been made to curb and to subdue these processes in order to obtain an artificial so-called cognitive discrimination which meets the severe demands of the arbitrarily established criteria.[4] Absence of these fundamental factors in judgment probably can be partially effected by those subjects only who have learned the art of focalizing their attention upon the more subtle aspects of sense data—to segregate from the tonal complex certain evasive attributes which are sometimes apparent to the professional analyst, and at the same time to inhibit completely all emotional attitudes which the sentienda might directly engender. Such undivided concentration is not a trait of childhood nor even of the average adult. Moreover, the interpretation of any results obtained through emotional repression and through the inhibition of other normal concomitant impulses becomes a vague uncertainty. As Malmberg and others have testified, it is the individual who is musically inclined, the individual who one might say is endowed with musical talent, that finds emotional repression and the inhibition of musical association so extremely difficult to effect. It is indeed difficult to conceive how a Brahms or a Schumann might restrict judgment of a harmonic clang to 'a pure cognitive act of discrimination' by totally repressing feeling and musical association.

Since, from practically every investigation conducted, musical agreeableness assumes such a basic position in

[4] Seashore (**17d**) incorporates the assertions previously made by Malmberg (**13b**), who writes: "Restfulness—a feeling of completeness, finality, or satisfaction, with its opposite disquietude—a feeling of incompleteness, needing to be resolved, was first adopted as a fifth criterion, but it soon developed that it must be dropped as it is a variable criterion directly due to progression and association, which must be excluded." Seashore (**17g**) further remarks that "in asking whether a two-clang is consonant of dissonant, therefore, we do not ask if it gives an agreeable or disagreeable feeling or which is the more agreeable or satisfying, but we require an æsthetic judgment in terms of the three specified criteria." As I have already maintained (**5d**), a simple assertion that, because of the test requirements, all feeling-tone is dispensed with in judgment is no valid proof of its actual absence.

consonance discrimination, we are forced to ask: *are* the proposed criteria—*blending*, *smoothness*, and *fusion*—sufficiently adequate in obtaining a measure of the sense of consonance? The answer to this question largely depends upon what one means by a 'sense of consonance.' To Malmberg, Gaw, and Seashore, the answer to this question is *yes*. To them, apparently, consonance is something quite *fixed*. They shun the introduction of musical agreeableness into the test because they rightfully recognize the relative nature of feeling-tone. Such relativity in response would seriously jeopardize the purpose for which the test was designed. Yet they do not seem to take into account that to the musician and to the individual musically inclined this relativity of feeling-tone is a fundamental and inseparable factor in nearly every judgment of consonance. It is quite needless to emphasize this point. The history of music has sufficiently convinced us. The sense of consonance (assuming the possibility of such a 'sense') can never be considered a fixed, absolute thing or experience, subject to the measure of some preconceived test-form. On the contrary, the sense of consonance is variable and progressive—progressive in the sense that the hedonistic factors involved find their source and development in the gradual adaptation of the end-organ of hearing to new complex tonal combinations. It is a well-established fact (**14**) that "the degree of consonance increases as a result of the frequency with which the interval is heard. Moreover, each interval, in its development from dissonance to a high degree of consonance, undergoes a characteristic affective development corresponding to the different stages in its course." What Liszt considered consonance would have met the sharp disapproval of Beethoven's listeners; and what Strawinsky and Honegger consider consonance would, in all likelihood, have shocked the musicians of the time of Liszt. The sense of consonance as a medium of appreciation in the life of the musician is not static. It is dynamic and has passed and continues to pass through an evolutionary process.[5] In the light of what we know concerning adapta-

[5] Moore (**14a**) has well summarized for us that "the history of each of our present consonances has apparently been a gradual development as dissonance, followed by a

bility as one of the outstanding characteristics of the human organism in its relation to environmental changes, we have reason to look forward to the time when the interval of the minor Second will become as agreeable to the layman as it now is to the members of the ultra-modern school of music. Hence we may look upon the Seashore Consonance Test in some measure as an attempt to moralize musical experience: consonance is defined and measured not as it *actually* exists in the experience of the musician, but rather as it *ought* to exist in the experience of the musician.

A chief source of error in the formulation of Seashore's Consonance Test seems to lie in the uncritical acceptance of the theoretic assumption that the consonance value of any interval remains constant irrespective of its position in the pitch series. This assumption, which has been derived from Stumpf's laws of tonal fusion (**19**), forms a basis for the belief that consonance is fixed and static. Now what Stumpf actually meant by consonance it is extremely difficult to say. From a study of his chapter (**19a**) on 'Stufen der Tonverschmelzung,' we are led to believe that grades of consonance are in reality grades of fusion. Consonance, then, is to be considered as degree of tonal fusion or the approximation in consciousness of tonal unity from tonal duality. This subtle distinction is purely psychological, and is derived and established by Stumpf from his own introspections of certain binary phenomena. After an analysis of his own auditory experiences, he attempted to verify statistically the nature of his experiences through a comparison of data gathered from other subjects. The type of

relatively short transition period, and later a gradual development as consonance. The typical case in which we may observe all the gradations is that of the Third. As a dissonance it was too harsh for any use at all, but afterwards was considered a useful interval if properly restricted. Then follows a transition during which some seem to have felt it as a consonance while others still considered it dissonant. Later comes its general admission as a parallel interval, and still later its use in the final chord of a cadence. Each stage had its characteristic pleasure value. The Seventh, altogether too harsh for the twelfth century ear, comes later to give much pleasure in the mildly dissonant chord of the Dominant Seventh, and the boldness of contemporary radicals augurs for it an eventual transition to the recent status of the Third, that of the most pleasant of consonances."

experimentation which Stumpf conducted and the results therefrom upon which he principally based his theoretical conclusions lose their importance when subjected to critical scrutiny. It is interesting to note that since individuals who were musically inclined or endowed *only occasionally* reported qualitative unity in the case of the Octave, Stumpf employed 'unmusical persons' in order to obtain his desired results. Regarding such an unusual and inexplicable test procedure, Lipps (**12**) in surprise has remarked:

> From this, I should necessarily conclude, in so far as I have up to this point grasped the concept of fusion, that fusion is therefore found factually in a more complete form and in a wider range with those who are unmusical than with those who are musical. According to this, too, if consonance and fusion are identical in meaning, the fact of consonance would necessarily be found in a more complete form and with a wider range in the case of unmusical persons.

If we are to understand with Stumpf that fusion means the qualitative unity that persists in a binary combination or a chord as the result of a loss of tonal individuality, and if facility of fusion, identical with degree of consonance, is relative to the tonal capacity for fusion (this capacity inherent in the tones themselves), then through the medium of normal auditory experience one is necessarily obliged to discount Stumpf's contention that the degrees of fusion are independent of the tonal region within the tonal range, of the relative or absolute intensity, of the minimal deviations of the number of vibrations from the ratio, and unaffected by the addition of a third or fourth tone.[6]

According to Stumpf then, in a binary combination, such as *c'b'* for example, the degree of fusion remains the same whether the upper tone is sounded louder than the

[6] Stumpf (**19b**) declares that "Der Verschmelzungsgrad ist unabhängig von der Ton-region. Der Verschmelzungsgrad ist auch unabhängig von der Stärke, und zwar sowol von der absoluten als relativen Stärke." And in regard to the relative intensity of two different tones sounded simultaneously, Stumpf unequivocally declares that "solange sie (*i.e.*, the two tones involved) unterscheidbar bleiben, kann ich eine Veränderung der Verschmelzungsgrades nicht bemerken. Beispielsweise wenn ich *c* und *g* zuerst gleichstark, dann *c* merklich stärker als *g* (oder umgekehrt) angebe."

lower tone or the lower tone is sounded louder than the upper tone. Moreover, the degree of fusion still remains unaltered when both tones are characterized by the same intensity. A shift of the interval to some more remote position in the pitch series will not even modify the degree of fusion. And finally, this same unfailing degree of fusion is retained despite the inclusion of a third or fourth extraneous tone. For Stumpf, the interval *CE* is identical in consonance with the interval *c'e'* whether the intensity attribute is relative or absolute. I have still to discover the musician with a normal ear who is willing to admit the sensory reality of such a remarkable phenomenon. And I have still to discover some naïve listener who is willing to admit the existence of identical degrees of fusion with the addition of a third or fourth tone. If by consonance Stumpf meant the simple fact of interval relationship, his laws would immediately become intelligible. However, we are emphatically told that consonance is not the simple fact of interval relationship, but rather the psychological phenomenon of qualitative homogeneity singularly intrinsic to a specific clang. When we consider the experimental studies (**8**, **10**) that have been conducted to illustrate in consonance perception the significance of pure tones, of overtones, of difference tones, and of summation tones, it is a point most difficult to conceive just how the degree of fusion remains unaffected through changes in pitch and intensity, either relative or absolute. I have previously insisted (**6**) that the psychological status of any musical grouping is altered when the tonal attributes of the grouping are altered.

If tonal quality is coincident with the sensation-form (**15**) assumed by the end-organ of hearing, and if such sensation-form [7] is dependent upon the quantitative relationships existing between extensity,[8] intensity, and protensity, then

[7] Sensation-form is here conceived as "the total sensorial effect upon the organism of any one uniform sensation, or of any one moment of a variable sensation. It is the unit of reaction resulting from the necessarily simultaneous presence of the three attributes and does not, strictly speaking, represent a synthesis, since we can only synthesize what has previously been separately present to consciousness" (**15a**).

[8] Outstanding musicians regard the pitch series as a linear, scaler or extensity series and not as a qualitative series.

it is inaccurate to refer to the identical quality of two tones or two combinations of tones which vary in their primary tonal attributes or, in other words, differ in their respective sensation-forms. Variability of sensation-form throughout the pitch series is due to the fact that the ratio between extensity and intensity is *not* constant.

That Malmberg, Gaw, and Seashore have championed Stumpf's contentions is clearly indicated in their discussions of the nature of the consonance test, and by the manner in which they have indiscriminately shifted intervals about in various keys.

There is, however, a practical advantage in adopting *fusion* as a *criterion* of consonance, despite the several confusions which center about this concept. Seashore employs the *phonograph* as the instrument of tonal production. If the criterion of *purity* as originally adopted by Malmberg and Seashore were retained, the phonograph would never serve as an adequate tonal control. The condition of tonal purity as a psychological phenomenon would differ considerably with the use of different phonographs possessing different acoustical properties. But since differences in resonance and in intensity do not alter the nature and degree of tonal fusion, it makes little difference whether an investigator employs an antique model or one of the recent audion electrolas, whether he fixes the disc rate at 78 revolutions or 65 revolutions, whether he employs loud steel needles or those of the soft fiber variety, or whether he tests each subject individually at a fixed distance from the instrument or a group of subjects distributed over a large auditorium. All of these discrepancies are well taken care of by Stumpf's laws of tonal fusion which allow for numerous inconsistencies not compatible with experimental procedure [9] in scientific psychology.

[9] In the 'Manual of Instructions' (18a), Seashore states that "loudness appropriate to the room may be varied by proper selection of metal needles," but he does not state what 'appropriate loudness' consists of, nor just how to effect a 'proper selection.' That constant experimental conditions are not observed is quite evident in another statement which refers to the adoption of an arbitrary disc rate. Most unusual is the instruction relating to the record of the sense of time in which the listener is

I have insisted in previous articles (**5**, **6**, **7**) that the phonograph with a variable resonance chamber is not a scientific instrument of tonal production and for this reason is not adequate in investigations which deal with the more subtle aspects of discriminatory diagnosis. If a particular phonograph were employed to establish certain norms in musical talent, and the very same phonograph under identically the same conditions were employed in further investigations of the same test material, the results thus obtained might be compared with the established norms. Or if it were possible to match the tonal qualities of several instruments, norm comparison would not be amiss. However, in making tests at the demonstration salon of one of the largest phonographic corporations in this country I discovered that the resonance chambers of several instruments representing the same structural model possessed markedly different tonal dispositions. Where different models are employed, the dispositional differences are much greater. The phonograph is satisfactory as an instrument of musical entertainment in cases only where gross, complex effects are desired and not where fine tonal distinctions or subtle gradations of intensity are desired.

Critical Summary

1. Seashore has not experimentally demonstrated his contention that the sense of consonance is a fundamental and essential capacity of the musical mind.

2. The assertion that the sense of consonance is an inborn capacity or trait is merely an assumption which has never been proved to be factual. No scientist has succeeded in demonstrating that complex musical talents are transmitted as unit characters of Mendelian inheritance.

3. Experimental data obtained by other investigators have shown that (1) the intercorrelations of the Seashore Tests of Musical Talent are very low; (2) that the coefficients of correlation are extraordinarily low (as low as 0.02 reported

expected to discriminate between very fine differences in interval length. In this most critical test for which definite norms have been established, the tester is informed that he is at liberty to choose a disc rate of *either* sixty *or* sixty-five.

by Malmberg and Brennen) between these tests, especially the consonance test, and musical training, musical environment, musical feeling, and musical ability in terms of performance proficiency adjudged by professional experts; and (3) that the coefficients of reliability for the consonance test are likewise very low.

4. The experimental results obtained by Dr. Larson cannot be compared with those of the writer since Dr. Larson preliminarily trained her listeners to react in a specific manner to specific stimulus patterns.

5. The criteria of consonance, which have been borrowed from the conflicting theories of Helmholtz, Wundt, Stumpf, and Krueger, are relative, indefinite, and intricate in meaning and demand further definition and explanation for clarity of comprehension.

6. Seashore regards the perception of consonance as static rather than dynamic. His point of view is not supported by the history of music nor by the results from outstanding experimental investigations on the relativity of consonance and dissonance perception. It is doubtful whether consonance perception considered as a simple cognitive act of discrimination delimited by the proposed criteria possesses any positive value in musical appreciation.

7. Practically every investigation that has been conducted on consonance discrimination with the Seashore Test has suffered the persistent invasion of feeling-tone and musical association. As Malmberg and others have testified, it is the individual who is musically inclined or endowed with musical talent that finds emotional repression and the inhibition of musical association most difficult to effect in consonance and dissonance perception.

8. If the perception of consonance as a cognitive act of discrimination becomes affectively colored, the Seashore test results will tend to be negative. This fact has been further demonstrated by Dr. Larson.

9. A chief source of error in the formulation of the consonance test lies in the uncritical acceptance of Stumpf's theoretic assumption that the consonance value of any

binary clang remains constant irrespective of its relative or absolute intensity or of its position in the pitch series.

10. The phonograph with the variable resonance chamber is not a scientific instrument of tonal production in the commercial application of musical testing.

11. Until the Seashore test of consonance admits of the more fundamental aspects of consonance perception which function in musical experience, its validity as a measure of musical talent will continue to be questioned. And until the same test is more thoroughly standardized in regard to presentation, it will be impossible to compare the results obtained by different investigators, to establish any reliable coefficients of correlation, or to refer test results to the pre-established norms.

BIBLIOGRAPHY

1. Brennen, F. The relation between musical capacity and performance, Psychol. Monog., 1926, **36**, 190–248. (a) 231.
2. Brown, A. W. The reliability and validity of the Seashore tests of musical talent, *J. Appl. Psychol.*, 1928, **12**, 468–476. (a) 472.
3. Dunlap, K. Elements of scientific psychology, St. Louis, C. V. Mosby Co., 1922. Also *cf.* his chapter 'The foundations of scientific psychology' in Mysticism, Freudianism, and scientific psychology, St. Louis, C. V. Mosby Co., 1920.
4. Gaw, E. A. A revision of the consonance test, Psychol. Monog., 1918, **25**, 134–147. (a) 140.
5. Heinlein, C. P. An experimental study of the Seashore consonance test, J. Exper. Psychol., 1925, **8**, 408–433. (a) 432; (b) 410; (c) 421; (d) 409.
6. Heinlein, C. P. A brief discussion of the nature and function of melodic configuration in tonal memory, with critical reference to the Seashore tonal memory test, *J. Genet. Psychol.*, 1928, **35**, 54 *ff.*
7. Heinlein, C. P. The affective characters of the major and minor modes in music, *J. Comp. Psychol.*, 1928, **8**, 101–142.
8. Helmholtz, H. L. F. On the sensations of tone, London, 1912.
9. Jennings, H. S. Heredity and environment, *Scient. Mo.*, Sept., 1924, 225 *ff.* Also *cf.* his Life and death, heredity and evolution, Boston, Richard G. Badger, 1920.
10. Krueger, F. Die Theorie der Konsonanz, *Psychol. Stud.*, 1910, **5**, 294 *ff.*
11. Larson, D. L. An experimental critique of the Seashore consonance test, Psychol. Monog., 1928, **38**, 49–81. (a) 62; (b) 54; (c) 54 *ff.*; (d) 57.
12. Lipps, T. Psychological studies: consonance and dissonance, *Psychology Classics, II.*, The Williams and Wilkins Co., Baltimore, 1926.
13. Malmberg, C. The perception of consonance and dissonance, Psychol. Monog., 1918, **25**, 93–133. (a) 127; (b) 108.
14. Moore, H. T. The genetic aspect of consonance and dissonance, Psychol. Monog., 1914, **17**, 62 *ff.* (a) 38 *ff.*

15. ORTMANN, O. The sensorial basis of music appreciation, *J. Comp. Psychol.*, 1922, **2**, 227–256. (a) 235.
16. PETERSON, J. A functional view of consonance, PSYCHOL. REV., 1925, **32**, 31 *ff.*
17. SEASHORE, C. E. The psychology of musical talent, New York, Silver, Burdett, and Co., 1919. (a) 7; (b) 144; (c) 155; (d) 148; (e) 151; (f) 158; (g) 149, 159.
18. SEASHORE, C. E. Manual of instructions and interpretations for measures of musical talent, New York, Columbia Graphophone Co., Reprint from 1919 ed., 12. (a) 2.
19. STUMPF, C. Tonpsychologie, Leipzig, 1, 1883; 2, 1890. (a) 2, 127–218; (b) 136 *ff.*
20. WUNDT, W. Physiologische Psychologie, Leipzig, 1910, 6. Aufl., Bd. 2.

[MS. received May 7, 1929]

DISCUSSION

STANDARDIZING THE GRADING OF LABORATORY REPORTS

Among the efforts at standardization in psychology today appears nothing pertaining to the grading of laboratory reports or books. The form for this, proposed below, has worked at the College of the City of New York and is flexible enough for use under other conditions. To aid in its understanding a brief account of the conditions at C. C. N. Y. follows.

The experimental work is two hours a week, concurrent with two hours of theory, in a general laboratory course. This brevity of time accounts for certain limitations and leeways in the plan described, not inherent in the plan itself. Written reports of fifteen experiments a semester are required (R) and about the same number are optional (O), with a proviso for substituting an O for an R in case of absence. Usually an O and an R alternate, and if so paired on the same topic the better student often will use both in his write-up. It is expedient to grade the books but three or four times a term—preferably at the end of the 3d, 6th, 10th and 15th week respectively. This is enough for the purposes in view if a foundation is laid and a manual used, especially where upper classmen predominate. Employed for such a basis is a printed folder, called a Laboratory Guide, whose four pages embody (1) a variety of General Instructions, (2) a conventional Report Outline, (3) an illustrative Sample Report and (4) a completed Sample Grade-Chart similar to the one below. The 'Parts' of the outline, explained and illustrated, are I. Purpose: II. Procedure: III. Results: IV. Discussion: V. References. Supplementing the Guide the first two meetings are devoted to demonstrational elaboration of the outline in application to a suitable experiment.

The chart can be barely mimeographed (but easily printed) on an ordinary typewriter sheet. It is intended to be self-explanatory and is found to be so after comments on it in the Guide and classroom. The Sample Chart and its explanations, carried by the Guide, is a real asset since, placed early in the student's hands, it serves both to inform and to satisfy him as to what he is 'up against' in the manner and the matter of rating; and is not so liable to loss as the single grade-sheet is when given out at the start. The latter reaches him on his first set of reports.

The comments, fuller and clearer, than those below tend as follows. Mention is made of the expediency, for grading and other purposes, of uniformity in the requirements for all students of the course; then in turn the bearings, for course or for scientific purposes, of each heading and its subheads.

1. *Presentability*, as covering the order and the neatness of the parts and of the whole report; the use of ink, of 8½ x 11 in. record and standard graph paper and of manila-folder cover with identification key in the upper right corner.
2. *Reliability*, as covering the accuracy of all facts and figures, quotations and references, etc.; the sincerity or earnestness of approach as seen from 'internal evidence'; the independence of the work presented or freedom from copying; the promptness of the response to each call for the reports; and the success shown in keeping the grading chart in place throughout the term.
3. *Quantity*, as covering the extent to which all 'R' and any 'O' experiments as a whole and their respective parts are adequately accounted for.
4. *Quality*, as covering the use and propriety of tables, calculations and graphs based on the individual, sectional and other scores provided in class; the observations made, both extrospections and introspections, separately; and the use of drawings or diagrams in accord with laboratory usage, and of psychological quotations and technical terms where natural and effective.
5. *Originality*, as covering discussion, used to include explanations and conclusions, comparisons and criticisms, implications and applications, history and prediction, etc.; clearness, used in reference to all routine and discussional statements; proportion, used in relation to the proper weighting, in length and adequacy of treatment, of the different parts and subparts of the report; suggestions, used to refer to any original or attempted contributions to psychological fact, theory or method; and technique, used to mean the effectiveness of assembling the whole with regard to the psychological, scientific and esthetic interests involved, rather than the experimental operations.[1]

[1] Where group experiments are not predominant this order of importance regarding arrangement-manipulation would be reversed, of course.

COLLEGE OF THE CITY OF NEW YORK

PSYCHOLOGY LABORATORY GRADING CHART

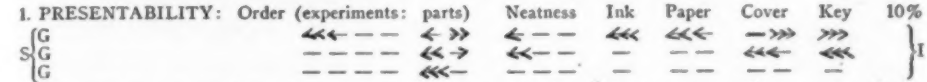
Keep this Chart as the first leaf of your Lab-book throughout the term.

The total credit is 100%. Each of the five main sources is credited roughly as shown by "%", save that repeated or unusual Superiority or Inferiority in any item may be scored more heavily than shown.

The work is Superior or Inferior in the items underscored, as the arrows point respectively towards S or I in the margin; increases in degree with increase in the number of heads used; is of doubtful or partial value if they point both ways; and has the value of the previous reading if none is used.

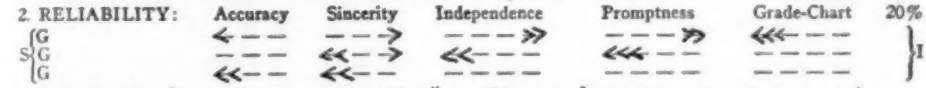
The "Gs" refer to successive Gradings of your work; the "Ns", to critical Notes for improving it, both here and in your reports; the "1—15", to the numbers of the "R" experiments; the "Os", to Optionals; and the "I—V", to the main parts of the Report Outline.

1. PRESENTABILITY: Order (experiments: parts) Neatness Ink Paper Cover Key 10%



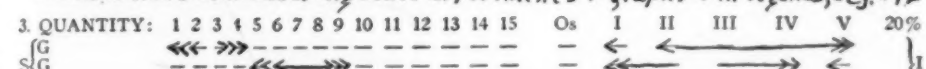
N Parts, not numbered; II-III & III-IV, mixed; book, coverless & keyless.
N Better, but show & name sub-parts of III & IV eachtime; see N on #7
N Mostly OK in above matters. Writing careless & numbers blurred.

2. RELIABILITY: Accuracy Sincerity Independence Promptness Grade-Chart 20%



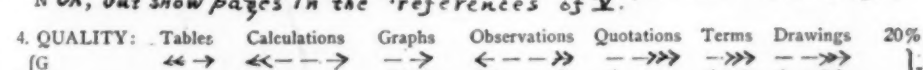
N Calculations wrong, per N in #3. Sincere? Independent? Late, 3 days.
N On sincerity, see N on #4. Scores & calculations wrong on #9.
N Inaccurate in numbering some experiments & graphs' & in legends, - cf. #12

3. QUANTITY: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 Os I II III IV V 20%



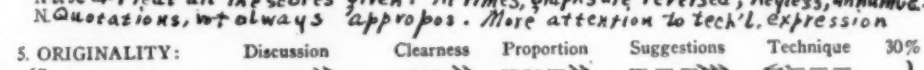
N Account for each Part. Add to III & IV. Show all scores given in class.
N Cite exact method, not merely that it was "oral", "printed" or "usual"; - cf. #8.
N OK, but show pages in the references of V.

4. QUALITY: Tables Calculations Graphs Observations Quotations Terms Drawings 20%



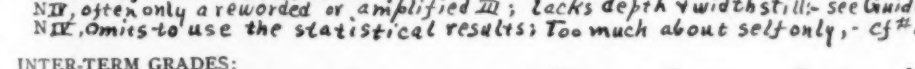
N Show table-footings. Consult graph rules, Introspections lacking. Quote.
N Use & treat all the scores given. At times, graphs are reversed, keyless, unnumb'd.
N Quotations, not always apropos. More attention to tech'l. expression

5. ORIGINALITY: Discussion Clearness Proportion Suggestions Technique 30%

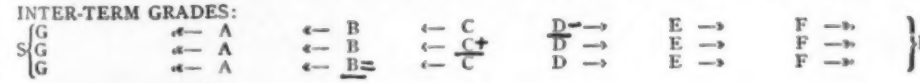


N III & IV often mixed; IV, too "easy" & restricted; see Guide. Use your results in interp'n.
N IV, often only a reworded or amplified III; lacks depth & width still; - see Guide.
N IV, omits to use the statistical results; Too much about self only, - cf. #15.

INTER-TERM GRADES:



FINAL TERM GRADE:



This sample charts a case, out of many different possible ones, as it would appear at the end of the term save that the arrowheads and notations would be made, more roughly, with colored pencil, a different color being usable for each presentation of work. *A* and *B* grade papers are apt to be more simple than this one because more satisfactory and uniform in their work, one underlining, with a single set of heads, often extending across the page and requiring nothing further there throughout the term. The given case shows the itemization rather than the generalization possible: as well known, the worse the student the more the trouble, for *any* method of handling. The successive subitems are ranked approximately in order of importance from left to right, and advantage of this is taken in any speeding up of the estimate of a lab. book's worth as well as in dealing with students who wish more information for improving their work or mark. These items run similarly from those always required, through those usually required, to those provisionally optional. It is thought that from this large array of data directly in view a fair impression of the mark sought will be had, not that any exact quantification of the items should be attempted.

The main advantages from the use of such a chart are that it

1. Summarizes, giving a compact bird's-eye view.
2. Illuminates the instructor as well as the student, acting for the former as a case-history for reference and for the latter as a personal record of achievement.
3. Keeps before the student his record as an incentive to improvement and allows him direct comparison with and help from other students.
4. Stimulates the student to analyze and to organize more clearly and more fully the aspects of experimental or scientific work in general and of that upon which he is engaged in particular.
5. Stimulates the instructor to standardize his own methods of working and grading and to itemize the account.
6. Renders possible the definite and the separate grading of every kind of item on the chart, or others, by the device of varying the number of heads.[2]
7. Dovetails with any further and fuller criticisms or notations (N) by the instructor or grader in the body of the lab. book.

[2] Greater precision may be attained if desired by making three arrowheads to the left mean 'A'; two mean 'B' and one, 'C'; and to the right, one mean 'D', two, mean 'E' and three, 'F.'

8. Is capable of expansion or contraction to meet different conditions—as to items covered, order ranked, per cents assigned, terminology used, etc.
9. Is flexible enough in its per cent-weight per unit and per total to allow for the instructor's subjective estimate of the student while retaining for the latter an impression of completeness and objectiveness in the appraisal.
10. Is adaptable to more than three or four markings per term; but would be less valuable where reports are called for frequently.
11. Is probably no more expensive of the grader's time than the usual method and after an initial adjustment period probably saves time.

The main disadvantages are (1) the grader's tendency to depend too solely upon the chart, omitting marginal notations *in situ* and (2) lack of ample space on a single page, to which it is best to limit it. Its ability to supplant the book notes often is an advantage, however—as where successive gradings need only identical notations.

In closing it may be said that hundreds of City College students have accepted this sort of marking, with a minimum proportion of misunderstanding or complaint. Nevertheless it is not thought to be the perfect pattern for the purpose devised.

H. D. MARSH

COLLEGE OF THE CITY OF NEW YORK

[MS. received March 16, 1929]

PURPOSIVE BEHAVIOR AND PREPOTENT STIMULUS

With reference to the criticisms by Tolman, Roberts, and Rosenow (November, 1928 issue of the REVIEW), concerning my paper on 'The Fundamental Error of the Concept of Purpose and the Trial and Error Fallacy' (PSYCHOL. REV., 1928, **35**, 414–433), Professor Weiss has come to my rescue (May, 1929 issue of the REVIEW) and has dealt with my critics so adequately that it seems unnecessary for me to write a rejoinder. However, there are still certain aspects of the discussion which must be clarified in order to make my position clearer.

Professor Tolman, after citing the experiments by Blodgett, Adams, and Elliott, says that the results of these experiments cannot be explained without the concept of Purpose. My answer would be that the simple idea of reenforcement of stimulus is

sufficient to take care of all these experimental facts. Is not the food (or the sign of food through conditioning) in the maze an additional stimulus reenforcing the organic stimuli from hunger contraction? In fact these experimenters, especially Blodgett, were really testing the effect of reenforcement of stimulus (food) on the behavior of the rat. But working under the delusion of trial and error learning, they all thought that they were experimenting on changes of the learning curve as effected by changes in food conditions. To me, if the so-called learning curve indicates anything at all, it indicates the inability or effectiveness of the experimenter to control stimulating conditions. *Certainly, the learning curve does not indicate animal learning: it simply measures the training of the trainer.* Why substitute so many verbal labels, purpose, drive, determining tendency, imitation, insight, ideation, gestalt or configuration, and what not, for a more careful analysis of stimulus? Why not have some more careful studies on the effect of summation, intensity, reenforcement, interference, and inhibition, localization, etc. of stimuli on behavior?

Professor Tolman says that in the present state of our knowledge of the stimulus-response connection, the concept of prepotent stimulus is not so helpful for immediate further experimentation as that of objective behavior-purposes. I do not think this is true. I believe that practically all of the experiments on animal learning in the past have been in the wrong direction and that we should start our animal work all over again, namely, that we should begin to investigate the fundamental properties (physical properties, I mean) of stimuli and their interrelations, so that we can have more definite control over behavior. The concepts of purpose, trial-and-error learning, imitation, insight, gestalt, etc. are great cloaks for our ignorance of such fundamental properties and are stumbling-blocks in the way of careful investigations and detailed analysis of stimuli-response connections. Of course, what Professor Tolman actually means to say is that no one can perform a trial-and-error learning experiment without the help of the concept of purpose. This is perfectly true (because the experimenter must read his own purpose into the animal beforehand). But the question I want to raise is: We have worked on trial-and-error learning for more than one quarter of a century but how much knowledge have we gained in controlling animal behavior? I should say, 'Very little.' The proper thing for us to do in the immediate future is to abandon all such experiments on trial-error learning,

together with experiments on imitation, ideation or insight, and gestalt, and have more intensive and extensive experiments on how to control stimuli so as to produce the desired responses of the animal in the quickest and surest way. Mere verbal concepts of purpose, insight, imitation, gestalt and the like, as Professor Weiss has already pointed out, do not help us to gain accuracy in predicting behavior.

Professor Weiss seems to have some objection to the concept of prepotent stimulus. In correspondence, he writes: "If you are going to place the behavior as mere series of movements, then you should also place the stimuli as a mere series of movements. To call a specific-grouping of stimuli prepotent is the same thing as calling a group of movements purposive. Stimuli just occur, that is all. They follow each other in a causal sequence (as you point out). Here again it is the experimenter who finds the prepotent stimulus. It is the experimenter who classifies the stimuli as prepotent or not, depending on whether the behavior of the organism that he is observing is being modified in the direction that he anticipates. If behavior is modified as he expects it to be he calls the group of stimuli which are doing it 'prepotent.' The stimuli themselves are merely light waves, sound waves, etc. They do not have an additional property of 'prepotency.'

"Can we not get along without the term prepotent? I think we can. As an experimenter I set out to modify the behavior of an organism. I vary the stimulating conditions until the behavior is modified in the direction set out beforehand. I can now describe this form of stimulation in terms of its physical properties. Is it necessary for me to add that this particular stimulus is 'prepotent'? Am I doing anything more than reiterating the fact that this particular stimulus brought about the changes in behavior? In biology, for instance, we do not say that a temperature of 80 degrees Fahrenheit is *prepotent* over 32 degrees so far as the rate of growth in plants is concerned. We merely announce the fact that plants grow faster under 80 degrees than under 32, and this statement is an adequate description."

I think I am in perfect agreement with Professor Weiss in this connection. When I used the word 'prepotent,' I did not mean to attach any particular significance or intrinsic value to it: I merely referred to the fact that in a given situation a group of stimuli are acting upon the organism, but as a result of physical intensity, summation, reenforcement, or inhibition, etc. Certain

stimuli, or a combination of several stimuli, become effective in calling out a response, while other stimuli at the moment may fail to do so. I called these effective stimuli 'prepotent' in that situation. *But they are 'prepotent' only when they are effective in eliciting a response in a given situation; in other situations they may become ineffective, and in that case, they are no longer prepotent. Prepotency is not a property of stimuli, nor can stimuli be classified according to their prepotency.* But now I think that it would gain much in clearness and cause less misunderstanding to drop the word 'prepotency' entirely, and describe effects of stimulating conditions purely in terms of intensity, summation, reenforcement, inhibition, etc. I am glad that Professor Weiss has pointed out the possible danger of the word 'prepotency,' and am taking this opportunity to publicly withdraw this word from the vocabulary of the science of behavior.

ZING YANG KUO

UNIVERSITY OF CHEKIANG, CHINA

[MS. received July 20, 1929]

NOTE ON USE OF THE TERM 'OBSERVER'

One of the words in common use in psychological literature that has become indiscriminately and inaccurately applied is the term 'observer.' Originally, the word was in use to designate the person whose experiences were the subject matter of observation, and who, in the nature of the case, had to serve as the first-hand observer whether he himself or another person be the 'experimenter.' 'Observer' and 'subject' were interchangeable terms. But by no means are all the problems of psychology concerned with a person's (the 'subject's') direct experience; and in the degree that they are problems of his efficiency, of his reactions or reaction tendencies, etc., they are a matter of observation less to him and more to the experimenter. In other words, in many contemporary lines of psychological investigation the so-called 'observer' does no observing!

One example from the literature will suffice. In a study of the effect of spectators upon the efficiency of an individual the subject was given the task of practicing an eye-hand pursuit coördination, first alone and later before an audience. Now, the data of the investigation included the objective quantitative record of his performance in each trial as noted by the experimenter, and also

reports by the spectators as to signs of confusion, uneasiness, *etc.*, but the subject himself was not called upon for report. Yet the subject was throughout the article referred to as the observer.

I propose that we drop entirely the misleading term 'observer' and use only its equally-respectable synonym, 'subject.' It should be noted that the latter term is used by the following compilers of experimental manuals: Sanford, Myers, Pièron, Langfeld and Allport, Foster, Hull, Woodrow, Poffenberger, Geissler, Kline and Kline, and Collins and Drever.

J. F. DASHIELL

UNIVERSITY OF NORTH CAROLINA

[MS. received August 24, 1929]

Psychological Review Publications

Original contributions and discussions intended for the Psychological Review should be addressed to

Professor Howard C. Warren, Editor Psychological Review,
Princeton University, Princeton, N. J.

Original contributions and discussions intended for the Journal of Experimental Psychology should be addressed to

Professor Madison Bentley, Editor Journal of Experimental Psychology,
Morrill Hall, Cornell University, Ithaca, N. Y.

Contributions intended for the Psychological Monographs should be addressed to

Professor Raymond Dodge, Editor Psychological Monographs,
Kent Hall, Yale University, New Haven, Conn.

Reviews of books and articles intended for the Psychological Bulletin, announcements and notes of current interest, and *books offered for review* should be sent to

Professor S. W. Fernberger, Editor Psychological Bulletin,
University of Pennsylvania, Philadelphia, Pa.

Titles and reprints intended for the Psychological Index should be sent to

Professor Walter S. Hunter, Editor Psychological Index,
Clark University, Worcester, Mass.

All business communications should be addressed to

Psychological Review Company, Princeton, N. J.

The Psychological Review

is indexed in the

International Index to Periodicals

to be found in most public and

college libraries

DIRECTORY OF
American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.; Cornell University. Subscription $6.50. 624 pages ann. Ed. by M. F. Washburn, K. M. Dallenbach, Madison Bentley and E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

The Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.; Clark University Press. Subscription $7.00. 700 pages ann. Ed. by Carl Murchison. Quarterly. Child behavior, differential and genetic psychology. Founded 1891.

Psychological Review—Princeton, N. J.; Psychological Review Company. Subscription $5.50. 500 pages annually. Bi-monthly. General. Founded 1894. Edited by Howard C. Warren.

Psychological Monographs—Princeton, N. J.; Psychological Review Company. Subscription $6.00 per vol. 500 pp. Founded 1895. Ed. by Shepherd I. Franz. Published without fixed dates, each issue one or more researches.

Psychological Index—Princeton, N. J.; Psychological Review Company. Subscription $4.00. 300–400 pp. Founded 1895. Edited by W. S. Hunter. An annual bibliography of psychological literature.

Journal of Philosophy—New York; 515 W. 116th Street. Subscription $4. 728 pages per volume. Founded 1904. Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush and H. W. Schneider.

Psychological Bulletin—Princeton, N. J.; Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Founded 1904. Edited by Samuel W. Fernberger.

Training School Bulletin—Vineland, N. J.; The Training School. Subscription $1. 160 pages ann. Ed. by E. R. Johnstone. Founded 1904. Monthly (10 numbers). Psychology and training of defectives.

Archives of Psychology—Columbia University P. O., New York City. Subscription $6. 500 pp. per vol. Founded 1906. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study.

Journal of Abnormal Psychology and Social Psychology—Eno Hall, Princeton, N. J. Subscription $5. Ed. by Morton Prince. In cooperation with Henry T. Moore. Quarterly. 432 pages ann. Founded 1906. Abnormal and social.

Psychological Clinic—Philadelphia; Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Founded 1907. Without fixed dates (9 numbers). Orthogenics, psychology, hygiene.

Psychoanalytic Review—Washington, D. C.; 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Founded 1913. Ed. by W. A. White and S. E. Jelliffe.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $6.00. Founded 1916. Bi-monthly. Ed. by Madison Bentley.

Journal of Applied Psychology—Baltimore, Md.; Williams & Wilkins Co. Subscription $4. 400 pages annually. Founded 1917. Quarterly. Edited by James P. Porter and William F. Book.

Journal of Comparative Psychology—Baltimore; Williams & Wilkins Company. Subscription $5. 500 pages annually. Founded 1921. Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.

Comparative Psychology Monographs—Baltimore; Williams & Wilkins Co. Subscription $5. 500 pages per volume. Edited by W. S. Hunter. Published without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Worcester, Mass.; Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Ed. by Carl Murchison. Monthly. Each number one complete research. Child behavior, differential and genetic psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J. Edited by W. S. Hunter. Subscription $6.00. Monthly. 700 pages annually. Founded 1927.

Journal of General Psychology—Worcester, Mass. Clark University Press. Subscription $7.00. 600–700 pages annually. Edited by Carl Murchison. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

PR.